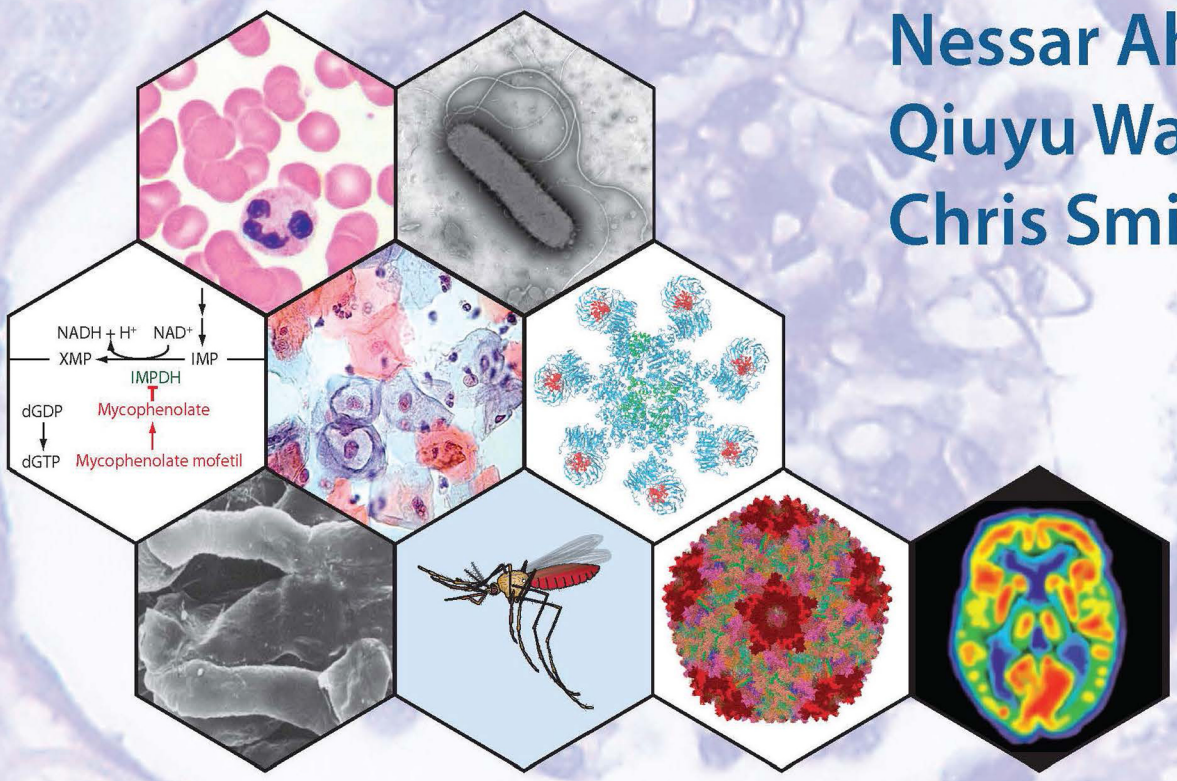
Second Edition

Biology of Disease



Nessar Ahmed
Qiuyu Wang
Chris Smith

SUPPORT MATERIAL

CRC Press
Taylor & Francis Group

Second Edition

Biology *of* Disease

Biology of Disease, Second Edition

Praise for the first edition:

'An excellent introduction to the subject of pathology, disease and laboratory diagnostics ... ideal for undergraduate students on biomedical science and related degrees.'

Alison Cottell for Biochemistry and Molecular Biology Education

'When you first pick up this book and peruse the pages it is immediately encouraging. The presentation and production values are very good, the diagrams and photographs look professional and it is a substantial volume, all of which suggests a worthwhile textbook. I recommend this book to students and tutors as a text for any course on the biology of disease.'

D.J. Cook for the British Journal of Biomedical Science

Biology of Disease describes the biology of many of the human disorders and diseases that are encountered in a clinical setting. Assuming no prior knowledge of the principles of diagnostic science, chapters follow a consistent format. Each opens with a set of learning objectives, an introduction to the topic and a description of the relevant biochemistry, molecular and cell biologies and physiology. How this biology is altered during the development of diseases and their diagnoses, management and treatment are then discussed. Real cases are used to illustrate the importance of biology in understanding the causes of disease, as well as in diagnosis and therapy. This second edition has been thoroughly revised and updated throughout and is now presented in full colour, with more case studies, questions and illustrations.

Key Features:

- Summarises the biological bases of common diseases
- Discusses real cases, placing basic science in a clinical context for the student
- Includes learning objectives, end-of-chapter summaries and questions for self-testing
- Is extensively illustrated in full colour with explanatory line diagrams and photographs
- Is associated with additional online resources for students and lecturers, available at www.routledge.com/9781032284361

The second edition of *Biology of Disease* is primarily aimed at meeting the needs of both undergraduate and postgraduate students of the biomedical sciences. However, it will also be of utility to those studying topics related to medicine, such as clinical sciences, clinical physiology and health sciences. It should also provide an effective reference for students beginning medical school. The text will also interest students of biological sciences with a feel for clinical-related aspects of their subject areas.

About the Authors:

Nessar Ahmed is a Reader in Clinical Biochemistry, Department of Life Sciences, Manchester Metropolitan University, Manchester, UK.

Qiuyu Wang is a Senior Lecturer in Biomedical Science, Department of Life Sciences, Manchester Metropolitan University, Manchester, UK.

Chris Smith was formerly a Senior Lecturer in Biochemistry, School of Biology, Chemistry and Health Science, Manchester Metropolitan University, Manchester, UK.

Second Edition

Biology *of* Disease

Nessar Ahmed
Qiuyu Wang
Chris Smith

CRC Press
Taylor & Francis Group
Boca Raton London New York

CRC Press is an imprint of the
Taylor & Francis Group, an **informa** business

Cover designed by Mick Hoult

Second edition published 2026
by CRC Press
2385 NW Executive Center Drive, Suite 320, Boca Raton, FL 33431

and by CRC Press
4 Park Square, Milton Park, Abingdon, Oxon, OX14 4RN

CRC Press is an imprint of Taylor & Francis Group, LLC

© 2026 Nessar Ahmed, Qiuyu Wang and Chris Smith

Reasonable efforts have been made to publish reliable data and information, but the author and publisher cannot assume responsibility for the validity of all materials or the consequences of their use. The authors and publishers have attempted to trace the copyright holders of all material reproduced in this publication and apologise to copyright holders if permission to publish in this form has not been obtained. If any copyright material has not been acknowledged please write and let us know so we may rectify it any future reprint.

Except as permitted under U.S. Copyright Law, no part of this book may be reprinted, reproduced, transmitted, or utilised in any form by any electronic, mechanical or other means, now known or hereafter invented, including photocopying, microfilming, and recording or in any information storage or retrieval system, without written permission from the publishers.

For permission to photocopy or use material electronically from this work, access www.copyright.com or contact the Copyright Clearance Center, Inc. (CCC), 222 Rosewood Drive, Danvers, MA 01923, 978-750-8400. For works that are not available on CCC please contact mpkbookspermissions@tandf.co.uk

Trademark notice: Product or corporate names may be trademarks or registered trademarks and are used only for identification and explanation without intent to infringe.

ISBN: 978-1-032-28437-8 (hbk)
ISBN: 978-1-032-28436-1 (pbk)
ISBN: 978-1-003-29683-6 (ebk)

DOI: 10.1201/9781003296836

Typeset in Utopia
by Evolution Design & Digital Ltd.

Access the Instructor and Student Resources: www.routledge.com/9781032284381

Dedication

In memory of our friend, mentor and extraordinary biochemical educator, Professor Ed Wood.

For my Mum and Dad.
Dr Nessar Ahmed

For our beautiful granddaughter, Lara.
Drs Qiuyu Wang and Chris Smith

For my family.
Dr Qiuyu Wang

Taylor & Francis
Taylor & Francis Group
http://taylorandfrancis.com

Contents

Preface to the second edition . . . . . *ix*
Glossary . . . . . *xi*

1. NATURE AND INVESTIGATIONS OF DISEASE 1
2. PATHOGENS AND PARASITES 23
3. PREVENTION, CONTROL AND TREATMENT OF DISEASES . . . 53
4. IMMUNE SYSTEM 99
5. DISORDERS OF THE IMMUNE SYSTEM 129
6. TRANSFUSION AND TRANSPLANTATION 159
7. DISORDERS OF THE ENDOCRINE SYSTEM 191
8. DISORDERS OF WATER, ELECTROLYTES AND URATE BALANCES 239
9. DISORDERS OF ACID–BASE BALANCE 263
10. DIET AND DISEASE 279
11. DISORDERS OF THE GASTROINTESTINAL TRACT, LIVER, GALLBLADDER AND PANCREAS 315
12. TOXICOLOGY 357
13. DISORDERS OF THE BLOOD 383
14. DISORDERS OF THE CARDIOVASCULAR SYSTEM 421

15. GENETIC DISEASES 453

16. MEMBRANE AND CELL-RELATED DISORDERS. 483

17. CANCER 513

18. AGEING AND DISEASE 569

Index 587

Preface to the second edition

It is 20-or-so years since the first edition of *Biology of Disease* was published: much has changed. Sadly, Professor Ed Wood of Leeds University is no longer with us to help write the second edition of *Biology of Disease*. Also, Dr Maureen (Mo) Dawson has retired from Manchester Metropolitan University (MMU) and, very sensibly, wanted to enjoy her retirement without the hassle of writing another textbook. Chris Smith has also retired from MMU but, less sensibly, agreed to help in producing the second edition of *Biology of Disease*. Dr Nessar Ahmed is still active in the study of the biomedical sciences, and the authorship has been considerably strengthened by Dr Qiuyu Wang, with her expertise in molecular and cancer biology. Over the last couple of decades, much has also expanded in the subject matter covered by *Biology of Disease*. The huge increases in knowledge, particularly of the underlying molecular biology of diseases and advances in so many treatments, have meant that, as in the first edition, although we have endeavoured to be as comprehensive as possible, we have had to limit the number of topics. The structure of this second edition follows that of the first as described in its preface and, once again, we have assumed a basic knowledge of biochemistry and cell and molecular biology on the part of the reader.

Like the first, this second edition is designed to meet the needs of undergraduate and postgraduate students studying topics related to medicine, for example biomedical sciences (medical technology in the USA), clinical sciences, clinical physiology, health sciences and nursing. However, in recent years, modules devoted to the biology of disease have become more prominent. Such modules are part of Bachelor of Science/Master of Science programmes in biomedical science, which are often accredited by professional bodies such as the Institute of Biomedical Science (IBMS). In these modules, disease is studied in an integrated manner, whereby basic science is linked to clinical practice, which is, of course, the underlying ethic of *Biology of Disease*. We also hope the text might be of interest to students of the biological sciences who have an interest in the clinical-related aspects of their subject areas.

During the writing of the first edition of *Biology of Disease*, the scientific community almost invariably replied positively to our requests for information and illustrative material. We are immensely grateful to the many academics and clinicians who maintained this wonderful eagerness to assist during the writing of the second edition and who generously supplied illustrations. These contributors are, naturally, cited in the relevant figure legends. We are especially grateful to our graphic artist, Mick Hoult (*Nature Photography*, www.naturalhistoryimages.co.uk), who, once again, not only drew but also redrew (and often re-redrew to the power *n*) the many figures, and also proved to be an expert at tracking down relevant illustrative material. We also thank Jiyao Wang, PhD, of the National Center for Biotechnology Information (USA) for his helpful advice on constructing molecular models. We are also grateful to Mo, who interrupted her retirement to read early drafts of Chapters 4, 5 and 6. Any remaining mistakes are ours (although we will try and blame her!). However, we are also grateful to the staff of Asthma UK, Cancer Research UK, Christie Hospital NHS Trust, Great Ormond Street Hospital, Manchester Blood Transfusion Service, Manchester Royal Infirmary, NHS Digital University of Manchester and several commercial firms, whose personnel did not wish to be individually named. Lastly, our thanks are offered to the many unsung workers who ensure that marvellous databases, such as the Centers for Disease Control and Prevention (USA), the

National Institute of Allergy and Infectious Diseases (USA) and the Protein Data Bank (USA), are freely available to peruse. The help we have received from all these people supports one of Ed Wood's favourite maxims, 'science is a cooperative venture'.

We would like to thank our reviewers for their comments regarding the first edition and their helpful suggestions on how to improve the second. However, we are, of course, responsible for any errors and omissions. We hope that you will write to us and tell us how we can still further improve the book for the benefit of future editions, which, hopefully, will not be after a gap of 20 years. Our final thanks are given to the team at CRC Press/Taylor & Francis, especially Dr Joanna Koster (Senior Publisher – Medicine & Life Science) and Jordan Wearing (Development Editor), for their constant encouragement, and our production editor, Linda Leggio.

Glossary

Accuracy The ability of a test for an **analyte** to give results comparable to the true value. See also **Precision**.

Achlorhydria A low level or the absence of secretion of acid in the stomach.

Acidosis A failure to maintain an appropriate acid–base balance in the blood that results in its pH being below the reference range. See also **Alkalosis**.

Acquired immunity The development of true immunity arising from a **specific immune response** to an **immunogen**.

Actinomycosis A fungal infection that may occur following injuries to oral tissues and causes inflammation.

Acute Diseases or clinical events that have a rapid, often severe, onset, but are of short duration. See also **Chronic**.

Acute intravascular haemolysis A clinical condition resulting from an incompatible blood transfusion, particularly involving the ABO blood group system.

Acute phase proteins A set of plasma proteins associated with an **acute phase response**, for example C-reactive protein and fibrinogen.

Acute phase response A systemic response involving changes in the composition of the blood, including an increased neutrophil count and an increase in the amount of the **acute phase proteins**, which occur in certain clinical conditions such as trauma, infections, necrosis, tumours and other inflammatory events.

Acute rejection The rejection of an **allograft** that occurs usually within a few weeks following the transplant.

Aetiology The cause(s) of a disease.

Affinity The strength of the binding between an immunogen (or antigen) and its complementary antibody. See also **Avidity**.

Agammaglobulinaemia (or hypogammaglobulinaemia) Clinical conditions in which the γ-class of **immunoglobulins** is deficient.

Ageing Most widely understood to be the decrease in the ability to survive on growing older.

Agglutination The clumping together of cells because of their interactions with specific **antibodies**.

Alarmins Proteins that stimulate immune cells to release type 2 cytokines, such as IL-4, IL-5 and IL-13, which promote the secretion of immunoglobulin E from B cells as a defence against **helminth parasites**.

Alkalosis A failure to maintain an appropriate acid–base balance in the blood resulting in its pH being above the reference range. See also **Acidosis**.

Alleles Different forms of the same **gene**.

Allergens Immunogens that cause an **allergy**.

Allergy A term often used to describe a type I hypersensitivity, which results in **inflammation** on exposure to an **allergen**.

Allogeneic Transplants that involve donors and recipients who are genetically nonidentical. See also **Autologous** and **Syngeneic**.

Allograft A transplant between two genetically different people. See also **Isograft** and **Xenogeneic transplant**.

Allotopic expression A technique whereby a mitochondrial gene is inserted into the nuclear genome where it is expressed to produce a product that is imported into the mitochondria where it alleviates the effects of a mtDNA mutation.

Alternative pathway Activation of complement proteins in the absence of antibody by cell wall components of bacteria and yeasts.

Amenorrhoea An absence or stopping of the menstrual cycle. See also **Galactorrhoea** and **Oligomenorrhoea**.

Amniocentesis A method of prenatally sampling biopsy material using a syringe to withdraw a sample of amniotic fluid containing cells genetically identical to those of the fetus and which can then be grown in tissue culture. See also **Chorionic villus sampling**.

Anaemia A reduced amount of haemoglobin in the blood because of insufficient erythrocytes or insufficient haemoglobin in them.

Analytes Substances that are measured to assist diagnosis and to monitor the treatment of disease.

Analytical sensitivity The smallest quantity or concentration of an **analyte** that can be detected by an analytical method. See also **Analytical specificity**.

Analytical specificity The ability of an analytical method to detect only the **analyte** in question. See also **Analytical sensitivity**.

Anaphylatoxins Fragments of activated complement proteins, such as C3a, C4a and C5a, which promote **inflammation**.

Aneuploidy A situation in which the number of **chromosomes** is fewer or greater than an exact multiple of the **haploid** number. See also **Euploidy** and P**olyploidy**.

Aneurysm A bulge that usually develops in weak areas of an arterial wall.

Angiogenesis The formation of new blood vessels.

Anisocytosis An abnormal blood condition characterised by excessive numbers of erythrocytes that have variable but abnormally large sizes. See also **Poikilocytosis**.

Anosmia The loss of the sense of smell.

Anoxia The absence of oxygen.

Antibiotics Drugs used to treat bacterial and other infectious diseases. Most antibiotics are derived from microorganisms.

Antibodies Glycoproteins produced by the immune system that bind to **antigens**. See also **Immunoglobulins**.

Antigen A macromolecule (protein, lipopolysaccharide) that provokes an immune response and binds specifically to an **antibody**.

Antigen-presenting cells Leukocytes that take up an **immunogen** and process it to form a peptide that can be recognised by T_H cells resulting in their activation.

Antinuclear factors Antibodies against nuclear components that occur in some **autoimmune diseases** such as rheumatoid arthritis and systemic lupus erythematosus. See also **Extractable nuclear antigens**.

Anuria The failure to pass urine. See also **Oliguria**.

Apheresis Isolation and concentration of components of the blood for clinical use. See also **Plasmapheresis**.

Apoptosis (also called cell suicide) A genetically programmed release of enzymes and other proteins that eventually bring about the death of the cell.

Apoptosome A multimeric protein complex that catalyses the degradation of a cell by **apoptosis**.

Arrhythmias Disturbances in the heart's rhythmic contractions.

Arteriosclerosis See **Atherosclerosis**.

Ascites An abnormal accumulation of fluid ('ascites fluid') in the abdomen.

Aspergilloma An infective mass of the fungus *Aspergillus fumigatus* growing within an old wound in the lungs.

Ataxia The inability to control muscular movements.

Atheroma A hard yellow plaque consisting of a necrotic (dead) core, rich in cholesterol and surrounded by fibrous tissue, that gradually builds up on the inside of medium-sized arteries. See also **Atherosclerosis**.

Atherosclerosis or **arteriosclerosis** The simultaneous development of an **atheroma** in an artery wall and its sclerosis (abnormal hardening or fibrosis).

Atopy An inherited tendency to develop type I hypersensitivities exemplified by asthma, rhinitis and eczema.

Atrioventricular node The node situated between the **atria** and the **ventricles** of the heart that delays the transmission of impulses from the **sinoatrial node** to allow the atria to contract completely and the ventricles to fill with as much blood as possible before **systole** occurs.

Atrium (plural atria) The upper chamber(s) of the heart. See also **Ventricle**.

Atrophy The decrease in mass or physiological activities of body parts due disease or a lack of nutrition.

Auscultation The use of a stethoscope to listen for distinctive diagnostic sounds within the body.

Autoimmune disease A situation in which the **immune system** mounts an immune response against '**self**' tissues.

Autologous stem cell transplants The use of a patient's own stem cells, which were harvested and stored prior to a treatment that destroys the patient's own stem cells, for example radiation or high-dose chemotherapy. See also **Allogeneic** and **Syngeneic**.

Autosomes The non-sex **chromosomes** that comprise 22 homologous pairs of the 46 chromosomes found in human **diploid** cells.

Avidity A measure of the ability of an **antibody** to bind to multiple binding sites on its complementary **immunogen** (or **antigen**). See also **Affinity**.

Bactericidal An **antibiotic** or other drug that kills its target bacteria.

Bacteriostatic An **antibiotic** or other drug that prevents its target bacteria replicating.

Balanced diet A diet that supplies adequate energy distributed appropriately between carbohydrates, lipids and proteins along with the necessary amounts of vitamins, minerals, trace elements, water and nondigestible fibre.

Basophil One of the three types of **polymorphonuclear leukocyte**. See also **Eosinophil** and **Neutrophil**.

Bioprosthetic valves Porcine and bovine heart valves that have been treated with glutaraldehyde to stiffen them and to mask their antigenic determinants. They are used to replace diseased human heart valves.

Biopsy The removal of a piece of tissue from a patient for clinical analysis.

Black Death The **pandemic** of bubonic plague that spread across Asia and Europe in medieval times, so called because of the black buboes that occurred on the skin of patients.

Bleeding time The period of time following wounding before the flow of blood stops. See also **Clotting time**.

B lymphocytes Types of **small lymphocytes** that develop fully in the bone marrow. On stimulation by an **immunogen**, they develop into **antibody**-producing cells.

Body mass index (BMI) A generally accepted way to assess the weight of a patient, where BMI equals the weight of the person in kilograms divided by the square of his/her height in metres. See also **Ideal body weight**.

Bone resorption The release of Ca^{2+} from bone.

Bradycardia A sustained, abnormally low heart rate. See also **tachycardia.**

Broad-spectrum antibiotics Antibiotics that are effective against both Gram-positive and Gram-negative bacteria.

Bronchiectasis A condition in which the bronchi and bronchioles are abnormally dilated.

Brush border The microscopic extensions called microvilli on the apical surfaces, that is, luminal sides, of the plasma membranes of enterocytes lining the intestine.

Bundle of His A group of modified cardiac muscle fibres, called **Purkinje fibres**, that carry electrical impulses in the heart.

C3 convertase A proteolytic enzyme consisting of the activated complement protein C4b2a that cleaves complement protein C3 into two fragments: C3 and the larger C3b molecule that is an opsonin. See also **Opsonisation**.

C5 convertase A proteolytic enzyme consisting of the complement proteins C3b and C3 convertase that are involved in the formation of a **membrane attack complex**.

Cachexia Severe and prolonged muscle wastage (weight loss) associated with several clinical conditions, including **cancer**.

Cancer A general term covering a number of diseases in which the growth of certain body cells becomes uncontrolled forming a tumour, which may be benign or **malignant**. See also **Neoplasm**.

Capsid The complex of proteins that form a shell enclosing the nucleic acid of a virus.

Capsomeres The protein components of a viral **capsid**. See also **Protomer**.

Carbuncle An amalgam of several abscesses or boils.

Carcinogens Environmental agents that cause cancers.

Carditis General inflammation of the heart.

Carriers (asymptomatic) Individuals who harbour a **pathogen** but who are symptom free.

Cell cycle The orderly sequence of biochemical and morphological events that constitute the growth phases of a reproducing cell. See also **Interphase**, **Mitosis** and **Meiosis**.

Cell-mediated immunity The production of cells that kill or recruit other cells to kill host cells infected by viruses or other intracellular parasitic microorganisms. See also **Humoral immunity**.

Centromeres The site at which members of a pair of sister **chromatids** are joined together.

Cerebrovascular accident See **Stroke**.

Chancre The small ulcer that develops at the site of a syphilitic infection and is the first diagnostic sign of the disease.

Chemokines Low-M_r proteins that function as chemoattractants in **chemotaxis**.

Chemotactic The ability of substances secreted by cells to attract other cells to the area.

Chemotaxis The migration of leukocytes along the concentration gradient of a **chemokine** (a subgroup of the cytokines) towards the chemokine-secreting cell or tissue.

Chloride shift The exchange of for Cl^- for a HCO_3^- across the plasma membrane of an erythrocyte.

Cholelithiasis The formation of gallstones.

Cholestasis The failure of bile to reach the small intestine.

Chorionic villus sampling A method of prenatally sampling biopsy material using a catheter inserted into the uterus to remove a small tissue sample of the fetal chorion. See also **Amniocentesis**.

Chromatid One of the pair constituting a condensed replicated **chromosome** present at **mitosis** or **meiosis**.

Chromosome A complex of proteins and single molecules of DNA (**chromatin**) that carries the genetic information in the form of **genes**. In **interphase**, they occur as extended structures but, during **mitosis** and **meiosis**, they are condensed to form microscopically visible structures.

Chronic Persistent diseases or clinical events that develop gradually over a relatively long time. See also **Acute**.

Chronic rejection Immunological rejection of a transplanted tissue/organ that occurs over months or years by a combination of **cell-mediated immunity** and **humoral immunity**.

Chylomicron Large **lipoprotein** complexes that transport triacylglycerols, phospholipids and lipid-soluble vitamins from the gastrointestinal tract to the lacteals of the lymphatic system.

Chyme The watery mixture of gastric juice and partially digested food released from the stomach into the small intestine.

Classical pathway of complement activation The pathway initiated when certain classes of **antibody** bind to an **immunogen**. See also **Alternative pathway**.

Clinical audit A process whereby practices and procedures involved in patient care are monitored and, if necessary, revised to provide a more effective, efficient and cost-effective service.

Clinical biochemistry The science concerned with investigating the biochemical changes associated with diseases.

Clinical genetics The science concerned with the identification and treatment of genetic abnormalities.

Clinical manifestations of disease These include the **signs** and **symptoms** associated with the disease.

Clinical sensitivity The incidence of positive results by a clinical test in individuals affected by the disease. See also **Clinical specificity**.

Clinical specificity The incidence of negative results by a clinical test in individuals free of the disease. See also **Clinical sensitivity**.

Clotting time The period required to seal a wound with a clot that prevents further blood loss. See also **Bleeding time**.

Codon The combination of three bases in a mRNA that code for a corresponding amino acid residue in a polypeptide.

Cofactor proteins Those that bind to **zymogens** and their protease products and increase the specificity and speed of the activation of the zymogen.

Colitis An **inflammation** of the colon.

Communicable diseases Diseases transmitted from a patient to another individual; an alternative term for **infectious diseases**. See also **Noncommunicable diseases**.

Compensation The biochemical and physiological mechanisms that in acid–base disorders attempt to return the pH of the blood to its reference range.

Congenital diseases Diseases present at birth that may or may not have been inherited.

Conjugate vaccines Those vaccines formed when a relatively poorly immunogenic material is attached to a protein that is much more immunogenic.

Contact allergies Reactions produced by a number of chemicals that directly affect the skin leading to a delayed **hypersensitivity**.

Contagious diseases A discrete category of **infectious diseases** that can be transmitted only by direct contact between individuals.

Control samples Samples that are identical in composition to test samples except that they contain known concentrations of the test **analyte**.

Corneal arcus Opaque fatty deposits around the periphery of the cornea that occur in **familial hypercholesterolaemia**.

C-reactive protein A particular protein whose appearance in the blood is often used as a marker of infection.

Cretinism A congenital condition in which a child suffers from intellectual disabilities, muscle weakness, short stature and neurological signs and is often mute, which occurs because of the failure of the thyroid gland to develop normally during embryonic growth.

Cytogenetics The microscopic study of **chromosomes**.

Cytokines A general name given to a large family of proteins that resemble hormones in that they are secreted by cells and stimulate activities in other cells after binding to receptors on their surfaces.

Cytotoxic cells Leukocytes that directly destroy infected cells in **cell-mediated immunity**.

Cytotoxic T lymphocytes Leukocytes that develop from T_C cells, which have been stimulated by an **immunogen**.

Debridement The surgical removal of infected or necrotic tissue.

Degenerative diseases Diseases involving the progressive loss of body tissues and impairment of their functions: usually associated with ageing.

Dehydration A condition that occurs when water loss from the body exceeds its intake. See also **Overhydrated**.

Diarrhoea The abnormal frequent passage of faeces that are larger in volume and more fluid than normal. It is not a disease but a symptom of an underlying condition.

Diastole Relaxation of the heart's **ventricles**.

Dietary reference values Guidance values of the amounts of nutrients that should be ingested, rather than exact recommendations.

Diploid The presence of pairs of **chromosomes** in a cell. Normal human somatic (body) cells contain 23 pairs of chromosomes. See also **Haploid**.

Dissection (with respect to arteries) The separation of the layers of the wall of an artery.

Dominant Genetically controlled factors that are expressed phenotypically in both the **homozygous** and the **heterozygous** conditions. See also **Recessive mutations**.

Dysmorphic Features that differ in appearance from those that are generally accepted as 'normal'.

Dyspnoea Often colloquially called 'shortness of breath'. It is caused by a narrowing of the airways that occurs in respiratory conditions, such as asthma and chronic obstructive pulmonary disease, and results in an intense tightening of the chest, difficulties in breathing and a feeling of suffocation.

Dystonia A neurological disorder associated with muscular spasms.

Electrocardiogram The record of the electrical potentials associated with the spread of depolarisation and repolarisation through the muscle mass of the heart during its beat cycle.

Emesis See **Vomiting**.

Endocardium The layer of smooth tissue lining the inside of the chambers of the heart.

Endocrine diseases Diseases that arise from the over- or underproduction of hormones or from resistance of a target tissue to a particular hormone.

Endogenous A general term meaning originating from within the body. See also **Exogenous**.

Endoscopy The use of a relatively long, thin, flexible tube to obtain biopsy material. An endoscope is similar but uses a camera and a light at the end of the tube to visually examine internal tissues and organs.

Envelope The lipid membrane that some types of viruses acquire as they leave their host cells.

Enzyme immunohistochemistry The use of antibodies labelled with an enzyme. The binding of the **antibody** to its **antigen** can be localised in cells as the enzyme converts a colourless substrate into an insoluble, coloured compound that can be seen under the microscope.

Eosinophil One of the three types of **polymorphonuclear leukocyte**. See also **Basophil** and **Neutrophil**.

Epidemiology The study of how diseases spread in populations in relation to their causal factors. See also **Aetiology**.

Epitopes Small regions of immunogenic molecules that are specifically recognised by components of the immune system.

Erythrocytes (also called red blood cells) Anucleate **formed elements** rich in haemoglobin that transport oxygen around the body.

Essential Nutrients that the body cannot make from other compounds and must therefore be supplied in the diet. See also **Essential fatty acids** and **Nonessential**.

Essential fatty acids Fatty acids that must be supplied in the diet, since humans lack the enzymes necessary to produce them.

Estimated average requirement The mean daily amount of nutrient (or energy) thought to be needed by a population. See also **Lower reference nutrient intake**, **Reference nutrient intake** and **Safe intake**.

Euploidy The situation in which the number of **chromosomes** is an exact multiple of the **haploid** number. See also **Aneuploidy** and **Polyploidy**.

Exogenous A general term meaning originating outside the body. See also **Endogenous**.

Exons Coding sequences of bases found within a gene that are separated by intervening noncoding sequences called **introns**.

Exophthalmos Protruding eyeballs.

Extended spectrum β-lactamases Enzymes produced by certain bacteria, for example *Klebsiella* spp. and *Escherichia coli*, that can hydrolyse extended spectrum cephalosporin antibiotics.

Extractable nuclear antigens Antigens that lead to the formation of autoantibodies in certain diseases such as rheumatoid arthritis and systemic lupus erythematosus. See also **Antinuclear factors**.

Extravascular haemolysis The uptake and subsequent destruction of antibody-coated erythrocytes in the spleen and liver. See also **Immune adherence**.

Fainting A common term for **syncope**.

Familial hypercholesterolaemia An autosomal genetically dominant condition associated with defective receptors on liver cells for low-density lipoprotein particles.

Favism The haemolytic crisis brought on by the consumption of fava or broad beans (*Vicia fava*) in persons with glucose-6-phosphate dehydrogenase deficiency.

Fetal alcohol syndrome A clinical condition in a neonate resulting from prolonged exposure to ethanol during fetal development.

Fistula An abnormal passage from a cavity or tube to another cavity or free surface.

Fluorosis The mottling of the teeth as they form in the jaws caused by drinking water containing fluoride in concentrations greater than 12 ppm.

Formed elements A term used to distinguish the erythrocytes, leukocytes and platelets of blood from the solutes found in **plasma**.

Fragile sites Particular regions of **chromosomes** that are susceptible to breakage when the cells are cultured in the absence of certain chemicals such as folic acid.

Fulminant cholera The massive leakage of intracellular water into the gastrointestinal tract and the subsequent **diarrhoea** following an infection with *Vibrio cholerae*.

Galactorrhoea An inappropriate secretion of milk. See also **Amenorrhoea** and **Oligomenorrhoea**.

Gene mutation theory of ageing This theory suggests that an accumulation of mutations during the course of a lifetime leads ultimately to tissue and organ malfunctions and eventually death.

Genes The fundamental units of heredity.

Genetic diseases Diseases that arise due to defects in **genes** or **chromosomes**.

Genetic imprinting A condition that depends upon whether paternal or maternal inheritance of a trait has occurred.

Genetics The branch of biology concerned with heredity, **genes** and DNA.

Genome The complete sequence of bases in the DNA molecules, that is, all the hereditary material, of an individual.

Genomics The study of the complete complement of bases in the DNA of an organism.

Genotype The genetic or hereditary constitution of an individual. The term can also be applied to any particular pair of alleles at specific loci on a **chromosome**. See also **Phenotype**.

Glossitis An inflammation of the tongue.

Glycation The chemical (nonenzymatic) reaction between the free carbonyl groups of reducing acyclic sugars and amino groups of a protein to form an unstable Schiff base that then rearranges to a more stable Amadori product.

Glycosuria The presence of glucose in urine that occurs when blood glucose concentrations exceed the renal threshold.

Graft versus host disease A condition associated with the transfusion of whole blood, blood products, bone marrow or solid tissue(s) that contain(s) residual lymphocytes to an immunodeficient patient. The **small lymphocytes** present in the donated material recognise **immunogens** of the recipient as foreign and mount an immune response resulting in the rejection of the donated material.

Gummas Tumour-like growths of a rubbery consistency that some patients develop in benign late syphilis.

Gynaecomastia The development of enlarged breasts in males.

Haemagglutination The agglutination of erythrocytes that occurs when they are cross-linked by antibodies to form visible aggregates.

Haematemesis The presence of blood in vomit.

Haematology The study of blood and its disorders.

Haemoglobinopathies Clinical conditions that result from mutations in the genes for globin molecules.

Haemolysis The rupture or lysis of **erythrocytes** and release of their haemoglobin.

Haemoptysis The coughing up of blood-containing fluid from the lungs.

Half-life ($t_{1/2}$) The time required for the concentration of a drug or **xenobiotic** in the blood to decline to half its original value.

Haploid (N) The presence of half the **diploid** number of a **chromosome**.

Hapten A small molecule that is unable to stimulate a specific immune response but becomes immunogenic if presented to the **immune system** in an appropriate context, such as by its attachment to a protein molecule.

Hayflick limit The maximum number of times a normal cell population can enter the cell cycle and divide before entering the senescence phase.

Health The state of physical, mental and social wellbeing and not merely the absence of disease.

Heart attack See **Myocardial infarction**.

Helminths The informal name for parasitic worms such as some types of nematodes, cestodes and trematodes.

Helper T lymphocytes See **T_H cells**.

Hernia The protrusion of an organ or tissue out of a body cavity in which it is normally found.

Heteroplasmy The condition in which a cell has a mixed population of mitochondria containing normal and mutant mitochondrial DNA. See also **Homoplasmy**.

Heterozygous The condition in which the two **alleles** of a **gene** are different. See also **Homozygous**.

Hirsutism The development of inappropriate increased body hair on the face, chest, upper back and abdomen in females.

Histopathology laboratories Laboratories concerned with the investigation of disease by examining specimens of cells and tissues.

Homeostasis The maintenance of a stable environment within the body. See also **Homeostatic diseases**.

Homeostatic diseases Diseases that arise when mechanisms for controlling homeostasis are disrupted or defective.

Homoplasmy The condition in which every mitochondrial DNA molecule in a cell has the same mutation. See also **Heteroplasmy**.

Homozygous The condition in which the two **alleles** of a **gene** are identical. See also **Heterozygous**.

Hordeolums ('styes') Infections of the eyelid generally involving the lid margins, eyelid glands or eyelash follicles.

Human herpes viruses (HHV-1–8) There are at least eight herpes viruses that can infect humans and cause clinical conditions.

Human lifespan The maximum age that can be attained by a human, which is approximately 120 years.

Humoral immunity The production of antibodies to protect against **pathogens** and **parasites**. See also **Cell-mediated immunity**.

Huntingtin A protein from a single mutated gene in patients with Huntington disease that accumulates in the nucleus rather than in the cytoplasm of cells.

Hybridomas (or hybrid myelomas) Antibody-producing cells made by fusing cultured **plasma cells** with cells derived from a myeloma (a **plasma cell** tumour) *in vivo*. The resulting cell is immortal and produces a **monoclonal antibody**.

Hyperacute rejection Failure of a graft within minutes or hours of transplantation once the tissue has become revascularised due to the presence of preformed **antibodies** against graft **antigens** present in the plasma of the recipient.

Hypercapnia A high partial pressure of CO_2 (PCO_2). See also **Hypocapnia**.

Hyperplasia A thymic abnormality characterised by the presence of excessive numbers of cells following an increased rate of mitosis.

Hyperplastic A clone of cells formed by the proliferation of a single cell following mutations induced in its genome by initiators and promoters.

Hypersensitivity See **Immunological hypersensitivity**.

Hypertension Sustained blood pressure higher than that regarded as normal.

Hypervitaminoses Clinical conditions caused by a toxic excess of a given vitamin. See also **Hypovitaminoses**.

Hypocapnia A low partial pressure of CO_2 (PCO_2). See also **Hypercapnia**.

Hypogammaglobulinaemia See **Agammaglobulinaemia**.

Hypogonadism A clinical condition in males in which the testes are small and soft, sperm production is deficient and testosterone secretion is decreased. In females, there is impaired ovarian function, oestrogen deficiency and abnormalities of the menstrual cycle.

Hypotension A sustained abnormally low blood pressure.

Hypotonia Poor muscle tone.

Hypovitaminoses Clinical conditions caused by an inadequate dietary intake of a vitamin, its impaired absorption or insufficient utilisation of an adequate intake. They can also result from increased dietary requirements, for example in pregnancy, without a corresponding increased intake or increased excretion of a vitamin. See also **Hypervitaminoses**.

Hypovolaemia A reduction in the volume of blood in the body.

Hypoxaemia A low partial pressure of O_2 (PO_2).

Hypoxia An inadequate supply of oxygen to body cells.

Iatrogenic diseases Clinical conditions that arise as a consequence of treatment.

Ideal body weight (IBW) A method of assessing the weight of patients. For males 5 feet tall, it is 106 pounds, and this increases by an additional 6 pounds for every inch over the height of 5 feet. For females, the IBW at 5 feet tall is 100 pounds and this increases by 5 extra pounds for each additional inch. See also **Body mass index**.

Idiogram A diagrammatic representation or interpretive drawing of the **chromosomes** based on the physical features as seen in a **karyogram**. See also **Karyotype**.

Idiopathic diseases Diseases of unknown cause(s).

Immune adherence The binding of **monocytes** and **macrophages** to **erythrocytes** that are coated with immunoglobulin G molecules. See also **Extravascular haemolysis**.

Immune system The set of organs, tissues, cells and molecules that protect the body from diseases caused by **pathogens** and **parasites**.

Immunoassay The use of **antibodies** to quantify the amount of an **analyte** in a clinical sample.

Immunodeficiency diseases Conditions that occur due to an inadequate **immune system**.

Immunofluorescence A type of **immunohistochemistry** in which the **antibody** is labelled with a fluorescent dye.

Immunogen Any molecule or organism that stimulates a specific immune response.

Immunoglobulins A collective name for the five classes of **antibodies**.

Immunohistochemical techniques The use of labelled **antibodies** to detect **antigens** on or in cells.

Immunohistochemistry An analytical technique in which a labelled antibody is used to specifically stain marker molecules on cell surfaces. See also **Immunofluorescence** and **Enzyme immunohistochemistry**.

Immunological diseases Clinical conditions that occur when the immune system causes damage to the body's own tissues or when it is deficient and unable to respond appropriately to **pathogens** and **parasites**.

Immunological hypersensitivity A state in which an immune response, often to something seemingly innocuous, results in tissue damage.

Immunological memory A molecular memory that allows a rapid protective response on second or subsequent contact with an **immunogen**.

Immunology The study of the **immune system** and how it works.

Immunology laboratories Those involved in the study of the **immune system** and its responses in both healthy and diseased states.

Immunotherapy Treatments that manipulate the **immune system** to improve its response against, for example, a **cancer**.

Incidence rate The number of new cases of a disease in a population occurring within a specified period of time.

Incubation period The time that must elapse before a disease becomes apparent following exposure to the aetiological agent.

Infarct A segment of tissue damaged by a disruption of its blood supply. See also **Infarction**.

Infarction The death (necrosis) of a section of tissue following a disruption to its blood supply. See also **Infarct**.

Infection The successful persistence and/or multiplication of a **pathogen** or **parasite** on or within the host.

Infectious diseases Diseases caused by **pathogens** and **parasites**, such as viruses and microorganisms (bacteria, fungi, protozoa) and **helminths**.

Infective dose A measure of virulence, being the minimum number of organisms or virus particles that can cause an **infectious disease**.

Inflammasome Cytosolic multiprotein complexes found in the cytosol that regulate the activation of caspase-1 and induce **inflammation** as a response to **pathogens** but also to certain patient-derived molecules.

Inflammation The array of responses to infection and tissue damage, such as localised pain, redness, swelling and heat.

Interleukins A large assembly of proteins produced by cells of the **immune system** that constitute a subgroup of the **cytokine** signalling molecules.

Interphase The period of the cell cycle between **mitoses** or **meioses**.

Intervention studies Studies aimed at changing diet and lifestyle to attempt to reduce the **incidence rate** of a disease associated with current activities.

Intrinsic factor A glycoprotein secreted by gastric parietal cells that bind vitamin B_{12} in the stomach and carries it to specific receptors on the mucosal surface of the ileum.

Intron See **Exon**.

Isograft A transplant of a piece of tissue from one site to another on the same patient, for example skin grafting in burn lesions. See also **Allograft** and **Xenogeneic transplant**.

Isohaemagglutinins Antibodies against blood group antigens that cause the **agglutination** of **erythrocytes**.

Karyogram The photographic representation of an individual's **chromosome** complement stained and arranged in a specific order. See also **Idiogram** and **Karyotype**.

Karyotype The characteristic number, size and shape of the **chromosome** complement of an individual or species. See also **Idiogram** and **Karyogram**.

Kernicterus Brain damage caused by the accumulation of bilirubin at concentrations of 200 μmol/L and above in the brain.

Ketonaemia An abnormal accumulation of **ketone bodies** in the blood.

Ketone bodies Acetoacetate, acetone and β-hydroxybutyrate.

Ketonuria The excretion of **ketone bodies** in the urine.

Korotkoff sounds The clear characteristic sounds that can be heard with a stethoscope due to the turbulence generated as the blood flows through a partially occluded artery.

Kussmaul respiration The deep sighing breathing (hyperventilation) associated with some types of **metabolic acid–base disorders**.

Kwashiorkor A protein-energy deficient nutritional condition that develops in children who have been abruptly weaned when a new sibling is born. Derived from one of the Kwa languages of Ghana, it means 'the one who is displaced'. See also **Marasmus**.

Large granular lymphocytes The **mononuclear leukocytes** involved in nonspecific immunity.

Latent period The time when an illness is developing but overt signs and symptoms are not apparent.

Lectins Glycoproteins that bind to the carbohydrate residues of molecules or cell membranes.

Lesions Structural or functional abnormalities.

Leukocyte One of several types of 'white blood cell', each with its own function.

Leukocytosis A **leukocyte** count in excess of 20×10^9/L compared with reference values of $4–11\times10^9$/L.

Leukodepletion The removal of **leukocytes** from blood by filtering it through leukocyte-specific filters that trap the leukocytes, but not the smaller **erythrocytes** or **platelets**, and reduces the leukocyte count to less than 5×10^6/L.

Leukopheresis The removal of **leukocytes** from blood with the resulting plasma and erythrocytes being returned to the blood donor. See also **Apheresis**, **Plasmapheresis** and **Plateletpheresis**.

Life expectancy The average length of life of individuals in a population.

Lipofuscins Yellow-brown pigments produced by degeneration of cell membranes and organelles, probably by the free radical peroxidation of membrane lipids.

Lipoprotein particles Complexes of cholesterol, triacylglycerols and proteins involved in the transport of these lipids by the blood. They occur as high-density lipoprotein, low-density lipoprotein and very low-density lipoprotein particles. **Chylomicrons** formed from fatty materials absorbed by the gastrointestinal tract are also lipoprotein particles.

Lower reference nutrient intake The amount of a nutrient sufficient for the requirements of 2.5% of a population. See also **Estimated average requirement**, **Reference nutrient intake** and **Safe intake**.

Lumen The interior space or channel of a hollow body organ.

Lymphocytes mononuclear leukocytes that occur in the blood as **large granular lymphocytes** and **small lymphocytes**.

Lymphoid stem cells Haemopoietic stem cells found in bone marrow that can divide to form the precursors of **small lymphocytes**.

Lymphoma A tumour that arises from lymphoid tissue, such as a lymph node.

Lymphopenia A reduced **lymphocyte** count.

Macrophages Immune cells that differentiate in solid tissues from immature **monocytes**.

Major histocompatibility complex A region of chromosome 6 that encodes the membrane proteins that present antigenic peptides (derived from **nonself** proteins) on the surfaces of **antigen-presenting cells** to **T_H cells/T lymphocytes**.

Malignant tumour See **Cancer**.

Mammography An X-ray examination of the breast.

Mantoux reaction A typical type IV or delayed-type hypersensitivity that occurs following vaccination against tuberculosis.

Marasmus A chronic disorder that develops over a period of months to years caused by an inadequate protein-energy intake. See also **Kwashiorkor**.

Mast cell A type of granule-containing **leukocyte** that resembles a **basophil** but that is found in solid tissues.

Medical microbiology The science concerned with the detection and identification of pathogenic viruses and microorganisms.

Megakaryocytes Large precursor cells that eventually fragment to form **platelets** in the blood.

Meiosis (reduction division) The type of cell division in which the daughter cells receive only one of each type of **chromosome** and which results in the formation of **haploid** gametes. See also **Mitosis**.

Membrane attack complex The large, cylindrical, hydrophobic structure constructed from a single complement protein molecule of C5b, C6, C7 and C8, together with several molecules of C9. It can insert into a cell plasma membrane leading to lysis of the cell.

Memory cells Small lymphocytes that do not differentiate following immunological stimulation but await the next exposure to the same **epitope** or **immunogen** and enable a faster and quantitatively greater response than the initial immunological defence. See also **Immunological memory**.

Mesentery The double-layered, fan-shaped portion of **peritoneum** that extends from the posterior abdominal wall and wraps around the small intestine and contains the blood and lymphatic vessels and the nerves that supply the small intestine. See also **Mesocolon**.

Mesocolon A structure that has similar structure and functions to the **mesentery** but which holds the large intestine.

Metabolic acid–base disorder A metabolic or renal dysfunction that results in an abnormal blood pH. See also **Mixed acid–base disorder** and **Respiratory acid–base disorder**.

Metastasis The detachment of cells from a **primary tumour** that spread and form **secondary tumours**.

Methaemoglobin Produced by the oxidation of ferrous (Fe^{2+}) iron to ferric (Fe^{2+}) iron in the haem groups of haemoglobin. It cannot bind or carry dioxygen.

Mitogens Substances that stimulate cells to divide. See also **Mitosis**.

Mitosis The type of cell division in which each of the daughter cells receives a full complement of **chromosomes**. See also **Meiosis**.

Mixed acid–base disorder A clinical condition in which a patient presents with more than one type of acid–base disorder. See also **Metabolic acid–base disorder** and **Respiratory acid–base disorder**.

Monoclonal antibody An antibody that recognises only a single specific **epitope** on an **antigen** and which is produced by a clone of identical plasma cells. See also **Polyclonal antibody**.

Monocyte A type of **mononuclear leukocyte** that circulates in the blood for about 72 hours before entering solid tissues and developing into a **macrophage**.

Mononuclear leukocytes One of the two main groups of **leukocytes**, which can be distinguished from the other **polymorphonuclear leukocytes** in having a more rounded nucleus. They are subdivided into **monocytes**, **large granular lymphocytes** and **small lymphocytes**.

Monoploidy The presence of only a **haploid** (that is, 23) set of **chromosomes** rather than the normal **diploid** complement of 46 in a human cell.

Morbidity The effects of a disease on the patient.

Mortality The possibility of a disease causing death; it is usually expressed as a percentage.

Mucins Glycoproteins of M_r up to 20×10^6 that bind large quantities of water giving them viscous, gel-like properties.

Multifactorial A condition caused by several factors.

Multiorgan dysfunction syndrome Stage III of sepsis, which is characterised by lactic acidosis, decreased blood pressure, **hypoxia** and **oliguria**.

Murmurs The characteristic sounds formed when abnormal heart structures cause a turbulent flow of blood.

Mutation A change in the **genotype**, which occurs as a result of the incorrect replication of DNA or from a random change to it caused by physical or chemical agents.

Myocardial infarction ('heart attack') A medical emergency that occurs when some or all of the blood supplied to the heart muscle through the coronary arteries is cut off.

Myocardium The cardiac muscle-rich wall of the heart.

Natural killer cells A type of **large granular lymphocyte** that kills virus-infected cells and some tumour cells nonspecifically.

Neoplasm (or neoplastic diseases) Often used synonymously with **cancer**; it means, literally, 'new tumour' or 'new mass'.

Nephrons The functional units of the kidney: each is composed of a Bowman's capsule, a glomerulus and a tubule.

Neutrophil One of the three types of **polymorphonuclear leukocyte**. See also **Basophil** and **Eosinophil**.

Noncommunicable diseases Diseases that cannot be transmitted from one patient to another individual. See also **Communicable diseases**.

Nondisjunction The failure of **chromosomes** to separate during metaphase in **mitosis** or at either of the metaphases in **meiosis**.

Nonessential A nutrient that the body can make from other metabolites and which does not need to be supplied in the diet. See also **Essential**.

Nonhaemolytic febrile transfusion reaction A clinical condition that can be caused by the transfusion of **leukocytes** that results in an adverse reaction. It may be caused by the reaction between antibodies and leukocyte antigens in the recipient, resulting in the lysis of donated leukocytes and the release of cytokines from them.

Nonself A term used in **immunology** to signify macromolecules or cells that are foreign to the body. See also **Self**.

Nonspecific defences The first line of immunological defence, including responses such as **inflammation** and the **acute phase response**.

Normal range See **Reference ranges**.

Nosocomial infections Those infections acquired in hospital.

Notifiable diseases Diseases for which clinicians are legally required to supply the details of all new cases to a central register.

Nuclear envelope Composed of the outer and inner nuclear membranes that separate the nucleus from the cytoplasm.

Nucleocytoplasmic transport The movement of materials between the nucleus and the cytoplasm of a cell.

Nutrition The study of food and nutrients and how they are used by the body.

Nutritional diseases Diseases that result from an inappropriate intake of nutrients.

Oedema An excessive accumulation of water in intercellular spaces.

Oligomenorrhoea Infrequent or irregular menstruation associated with a number of clinical disorders. See also **Amenorrhoea** and **Galactorrhoea**.

Oliguria Passing less than 400 mL of urine per day. See also **Anuria**.

Oncogenes One of the two major groups of genes linked to **cancer**: they are mutated forms of normal genes called proto-oncogenes. See also **Tumour suppressor genes**.

Oncology The branch of medicine involved with the study of **malignant tumours**, including their epidemiology, diagnosis and treatment.

Opisthotonos An involuntary arching of the head, neck and spine backwards.

Opportunistic infections Infections that occur when a decline in the immunological functions of patients makes them susceptible to infections by viruses, microorganisms, yeasts and **helminths** that are normally nonpathogenic.

Opportunistic pathogen A virus, microorganism or **helminth** that causes **opportunistic infections**.

Opsonisation The process by which phagocytic cells bind readily to bacterial cells coated with the activated complement protein C3b, promoting phagocytosis of the cells.

Osmotic diuresis The production of large volumes of urine that occurs when blood glucose concentrations exceed the renal threshold and so glucose is lost in the urine.

Ototoxicity The loss of hearing or balance and/or the presence of tinnitus or dizziness caused by drug or chemical damage to the inner ear.

Overhydrated The accumulation of water in body compartments. See also **Dehydrated**.

Pacemaker See **Sinoatrial node**.

Palpitations Abnormal or irregular heartbeats of such severity that the patient is conscious of them.

Pandemic A disease epidemic that affects patients over a wide geographical area.

Papules Solid raised areas of skin about 5 mm in diameter whose shape and colour may vary.

Paraesthesia Abnormal sensations such as itching, prickling, pins and needles anywhere on the body.

Paraparesis A progressive weakening of the legs.

Parasite An organism that lives at the expense of another: often applied to viruses and some protozoa and **helminths**. The terms **pathogen** and parasite are virtually interchangeable.

Pathogen-associated molecular pattern A molecular signature, for example a particular pattern of sugar hydroxyl groups, that can be recognised by **pathogens** and **parasites**. See also **Pattern recognition receptor**.

Pathogenesis The molecular and biochemical mechanisms that allow **pathogens** and **parasites** to cause diseases.

Pathogens Viruses and organisms that cause disease. See also **Parasite**.

Pathology The scientific study of diseases.

Pattern recognition receptor A molecule that is complementary to a **pathogen-associated molecular pattern** and able to specifically bind to it.

Penicillin An antibiotic first discovered (in the mould *Penicillium notatum*) in 1928.

Pericarditis An inflammation of the **pericardium**; can be **acute** or **chronic**.

Pericardium The thin, flexible, two-layered bag surrounding the heart.

Peristalsis The rhythmic waves of smooth muscle contractions that propel food along the gastrointestinal tract.

Peritoneum The strong, transparent membrane with a smooth surface that lines the abdomen and binds the gastrointestinal tract and accessory organs to each other and to the inner wall of the abdomen.

Petechial haemorrhages Small, rounded, dark red spots caused by bleeding under the skin.

Phagotherapy The clinical use of bacteriophages to kill infectious bacteria.

Pharmacodynamics Describes how a drug interacts with its target site and if the resulting biochemical and physiological processes produce any therapeutic or toxic effects. See also **Pharmacokinetics**.

Pharmacokinetics Describes the uptake, distribution, metabolism and excretion of a drug or **xenobiotic** from the body. See also **Pharmacodynamics**.

Pharmacology The study of the effects of drugs in the prevention, diagnosis and treatment or cure of disease.

Pharmacotherapeutics The branch of pharmacology concerned with the administration of drugs to prevent, monitor and treat diseases.

Phenotype The visible or measurable characteristics of an individual or, indeed, any observable biological trait. See also **Genotype**.

Photoageing The characteristic fine and coarse wrinkling, blotchy pigmentation and rough skin texture induced by ultraviolet light from the sun that is a major factor contributing to the premature ageing of skin.

Phototherapy The use of light in treating clinical conditions, for example the repeated exposure for 36 hours of full-spectrum bright light to treat **Seasonal affective disorder**.

Plague A term originally applied to any widespread disease causing great **mortality** but now confined to bubonic plague, that is, an infectious disease of humans and some animals caused by the bacterium *Yersinia pestis*.

Plasma The ion- and protein-rich fluid remaining when **erythrocytes**, **leukocytes** and **platelets** have been removed from blood, but which retains an intact clotting system. See also **Serum**.

Plasma cells Antibody-secreting cells that develop from **B lymphocytes** under the influence of cytokines released from **T_H cells**.

Plasmapheresis The collection of **plasma**, with the return of the **erythrocytes** and **leukocytes** to the blood donor. See also **Apheresis**, **Leukopheresis** and **Plateletpheresis**.

Platelet aggregation Occurs upon wounding when **platelets** extend pseudopodia and become sticky and when stimulated by ADP clump together to form a plug that prevents further blood loss.

Plateletpheresis The collection of **platelets** from blood with all other components being returned to the blood donor. See also **Apheresis**, **Leukopheresis** and **Plasmapheresis**.

Platelets Vesicle-like, subcellular fragments about 3 μm in diameter with a volume of 7 fL (femtolitres) that have clotting functions and are formed by the fragmentation of **megakaryocytes**.

Poikilocytosis An abnormal variation in the shapes of erythrocytes. See also **Anisocytosis**.

Polyallelic The occurrence of many different **alleles** of a **gene**, although any one individual can only posses a maximum of two alleles at each locus, given that **chromosomes** occur in pairs.

Polyclonal antibodies Each mature, differentiated B lymphocyte or **plasma cell** produces a single type of antibody. However, the response of the immune system to a single **immunogen** yields hundreds of different clones of B cells so that the response of the system is polyclonal, resulting in the appearance of a heterogeneous array of **antibodies** in the blood. See also **Monoclonal antibody**.

Polydipsia Intense thirst.

Polymerase chain reaction An experimental, *in vitro* method for multiplying short strands of specific DNA fragments in an exponential manner.

Polymorphonuclear leukocyte One of the two main types of **leukocytes** that can be distinguished from the other type (**mononuclear leukocytes**) in having lobed nuclei and granular cytoplasm.

Polymorphonucleocyte See **Polymorphonuclear leukocyte**.

Polyploidy The situation in which the **chromosome** number is an exact multiple of the haploid number but exceeds the diploid number. See also **Aneuploidy** and **Euploidy**.

Polyps Benign tumours that arise, for example in the nose or from the epithelium of the gastrointestinal tract.

Polyuria The production of abnormally large volumes of urine. See also **Anuria** and **Oliguria**.

Portals of entry Sites at which **pathogens** and **parasites** can enter the body.

Portals of exit Sites at which **pathogens** and **parasites** can leave the body.

Precision The ability of a test for an **analyte** to give the same result every time the same sample is analysed. See also **Accuracy**.

Predisposing factors Conditions or situations that make an individual more susceptible to disease; they include age, sex, heredity and environmental factors.

Premalignant Clinical conditions that increase the risk of the patient developing a **malignant tumour**.

Prevalence The proportion of people in a population affected by a disease at a specific time.

Primary immunodeficiency diseases Clinical conditions that result from a genetic defect leading to failure of one or more components of the immune system. See also **Secondary immunodeficiency diseases**.

Primary tumour Malignant tissue growing at the site where it initially started. See also **Metastasis** and **Secondary tumours**.

Prognosis The likely outcome of a disease.

Programmed theory of ageing Suggests that each species has an inbuilt biological clock and that ageing is a genetically programmed series of events.

Proinflammatory Description of a number of **cytokines**, which when secreted promote **inflammation**.

Prophylactic bilateral mastectomy The removal of apparently healthy breasts to prevent the development of breast **cancer** in genetically susceptible women.

Proteasomes Multisubunit structures that degrade (hydrolyse) proteins tagged with ubiquitin in an ATP-dependent manner.

Protomer The protein components of viral **capsomeres**.

Psychogenic diseases Clinical conditions that originate in the mind.

Ptosis An inability to raise the eyelids.

Pulmonary circulation The path that the blood follows between the right side of the heart, the lungs and the left **atrium**. See also **Systemic circulation**.

Pulse The bulging of the arteries that can be felt at several sites in the body during **systole**.

Purified protein derivative A mycobacterial extract used in the **Mantoux reaction** to test if an individual is already sensitised to tuberculosis by infection or vaccination.

Purkinje fibres See **Bundle of His**.

Pyogenic infections Bacterial infections that cause the formation of pus.

Recessive mutations Mutations that are only expressed in an individual who is **homozygous** for the condition.

Recommended daily amount Originally used by the UK Department of Health in 1979 to define the amounts of certain nutrients needed by different groups within the population.

Reference nutrient intake An amount of a nutrient that is greater than the dietary needs of 97.5% of a population. See also **Estimated average requirement**, **Lower reference nutrient intake** and **Safe intake**.

Reference ranges Numerical reference limits based on the mean value of an **analyte** plus or minus two standard deviations against which test results can be compared. The term **normal range** is still used synonymously with reference range.

Reflux oesophagitis A backflow of gastric juice from the stomach into the lower end of the oesophagus.

Regurgitation An imperfect closing of the heart valves leading to leakage and backflow of blood. See also **Stenosis**.

Relapse Occurs when, following a period of apparent **remission**, the **signs** and **symptoms** of the disease return.

Remission A period of good health with a reduction or disappearance of symptoms following a disease; may be permanent or temporary. See also **Relapse**.

Respiratory acid–base disorder A lung dysfunction that results in an abnormal blood pH. See also **Metabolic acid–base disorder** and **Mixed acid–base disorder**.

Resting membrane potential The potential difference across the plasma membrane of about –60 mV, which occurs as a result of the activity of its membrane-bound Na^+/K^+-ATPase.

Reticulocytes The immediate precursors of **erythrocytes**, which have nuclei and are still able to synthesise haemoglobin.

Retinoids The name for a group of naturally occurring, as well as many synthetic, forms of vitamin A (retinol).

Reversion Occurs in some patients who do not develop an effective immune response to an attenuated vaccine and the weakened components in it become virulent.

Rheumatoid factors Autoantibodies present in the plasma of patients with rheumatoid arthritis that seem to be directed against the patient's own immunoglobulin G molecules.

Risk factors Certain dietary, occupational or lifestyle conditions that increase the chances of developing a disease.

Safe intake The amount of a nutrient judged to be sufficient for the needs of most people. See also **Estimated average requirement**, **Lower reference nutrient intake** and **Reference nutrient intake**.

Screening The use of laboratory tests to detect a disease before it presents clinically.

Seasonal affective disorder A condition characterised by changes of mood and of eating and sleeping patterns associated with an inappropriate secretion of melatonin in winter.

Secondary A term often used to refer to a condition that arises from an existing disease. It is also used to describe a stage of a disease.

Secondary immunodeficiency diseases Arise as a consequence of other conditions. See also **Primary immunodeficiency diseases**.

Secondary tumours Formed when a primary tumour undergoes **metastasis**, spreading to other tissues and organs. Often loosely referred to as 'secondaries'.

Secretors Individuals who express the A, B and H blood antigens on soluble proteins and secrete them into saliva and other body fluids.

Secretory component A protein that protects dimeric immunoglobulin A molecules from enzyme-catalysed hydrolysis.

Self A term used in **immunology** to refer to the components that make up the body itself. See also **Nonself**.

Senescence The decline in the functions of almost all parts of the body and at all levels of organisation, from cells to organ systems, that occurs during **ageing**.

Sepsis The presence of **pathogens** in the blood of a patient.

Serous fluid A liquid that resembles serum found, for example, in the peritoneal cavity. See also **Ascites**.

Serum The clear straw-coloured liquid remaining after the blood has been allowed to clot and all the **formed elements** have been removed.

Severe combined immunodeficiency A life-threatening clinical condition involving deficient numbers of both B and T lymphocytes.

Shock A life-threatening condition that occurs if the blood pressure of the patient suddenly falls too low.

Sickle cell anaemia A haemoglobinopathy characterised by the presence of crescent-shaped **erythrocytes**.

Sign Something a clinician specifically looks or feels for when examining a patient.

Sinoatrial node The natural 'pacemaker' of the heart, which generates the electrical impulses that regulate the rate of heart contractions.

Sinusoids Blood vessels that occur instead of capillaries and venules in organs such as the liver, spleen and bone marrow. They are lined with a discontinuous reticuloendothelium that allows blood components to pass between the vessels and surrounding tissue.

Small lymphocytes The **mononuclear leukocytes** involved in the **specific immune response**.

Sonogram A picture of body tissues/organs produced by **ultrasonography**.

Specific immune response The biochemical and physiological mechanisms that allow the development of true specific immunity against an infectious agent.

Spirometry Diagnostic test(s) to assess pulmonary functions by measuring lung volumes to investigate respiratory conditions, such as allergic asthma and chronic obstructive pulmonary disease.

Splicing A posttranscriptional change in which **introns** are excised from a nascent RNA transcript and the newly formed ends of the **exons** are joined together.

Stem cell transplant A transplant of stem cells using bone marrow, peripheral blood or umbilical cord as the source of the cells.

Stenosis A failure of the heart valves to open appropriately. See also **Murmurs** and **Regurgitation**.

Strabismus A squint or abnormal alignment of the eyes.

Stroke (or cerebrovascular accident) The death or permanent damage of brain tissue that occurs if the blood supply to that part of the organ is disrupted for sufficient time and the brain cells are permanently damaged or die due to lack of oxygen.

Subclinical A stage at which a disease is established even though overt **signs** and **symptoms** are not apparent.

Superantigens Toxins secreted by bacteria such as *Streptococcus pyogenes* or *Staphylococcus aureus* that stimulate the release of cytokines from immune cells and produce an excessive inflammatory response called **systemic inflammatory response syndrome**.

Symptom An indication that a disease is present and something that the patient complains about, for example nausea, malaise or pain.

Syncope (faint) A loss of consciousness resulting from a temporarily inadequate supply of oxygen and nutrients to the brain.

Syndrome A term applied to describe certain diseases that are characterised by multiple abnormalities and features that form a distinct clinical picture.

Syngeneic stem cell transplants Involve donors and recipients who are genetically identical, for example between identical twins. See also **Allogeneic** and **Autologous stem cell transplants**.

Systemic circulation The flow of blood from the left ventricle to all the tissues of the body with the exception of the lungs and back to the right atrium. See also **Pulmonary circulation**.

Systemic disease A disease that affects the whole body.

Systemic inflammatory response syndrome See **Superantigens**.

Systole Contraction of the heart ventricles. See also **Diastole**.

Tachycardia A sustained and inappropriately fast heartbeat. See also **bradycardia**.

T_C cells (cytotoxic T lymphocytes) Leukocytes capable of killing virus-infected cells.

Thalassaemias A large group of related haemoglobinopathies of varying severity that originate from point mutations or deletions in globin genes.

T_H cells (helper T lymphocytes) Leukocytes that develop into cytokine-secreting T_H cells when stimulated by an **immunogen** and which produce an array of cytokines that control the activities of cells of the immune system.

Tissue tropism The property displayed by a virus in being able to infect only a specific type of host cell.

T lymphocytes Leukocytes derived from **small lymphocytes** that leave the bone marrow in an immature state and, in the fetal thymus, mature into **T_C** or **T_H cells**.

Toxic diseases Diseases caused by the ingestion of a variety of poisons that may be encountered in the environment.

Toxicology The study of the adverse effects of toxic chemicals on organisms.

Toxoids Inactive bacterial toxins that are used in the production of vaccines. See **Vaccination**.

Transgenic organisms Organisms that express genes from a different species that have been inserted into their genome.

Transplant The donation to a patient of solid tissues, organs or haematopoietic stem cells harvested from blood or bone marrow.

Traumatic diseases Diseases caused by physical injury, including mechanical trauma.

Tumour An abnormal mass of cells that may be benign or malignant.

Tumour suppressor genes One of two major groups of genes whose mutated forms are linked to the formation of cancers. See also **Oncogenes**.

Ultrasonography A technique that uses the reflections of ultrasound waves from internal tissues or organs to form a **sonogram**.

Uraemia The presence of high concentrations of urea in the blood.

Uricolysis The degradation of urate (uric acid) to CO_2 and NH_3.

Urticaria Transient skin rashes characterised by wheals of varying shapes and sizes that have pale centres surrounded by red margins.

Vaccination The process of protecting people from infection by deliberately exposing them to **pathogen** or **parasite** components to initiate a protective immune response.

Vegetation An accumulation of bacteria and blood clots on the valves of the heart.

Ventricle Lower chamber of the heart.

Virilism The development in females of inappropriate male characteristics, for example **hirsutism**, increased muscle mass, deepening of the voice and male psychological characteristics.

Virulence factors Features of **pathogens** that allow them to cause **infectious diseases**.

Vitamers A term applied to a group of closely related compounds that all possess the activity of the named vitamin.

Vomiting (or emesis) The forced expulsion of food from the stomach (and sometimes duodenum) through the oesophagus and mouth.

***WAS*-related disorders** The clinical conditions Wiskott–Aldrich syndrome, X-linked thrombocytopenia and X-linked congenital neutropenia, all of which occur because of more than 300 mutations in *WAS*.

Winter depression See **Seasonal affective disorder**.

Xanthomas Yellowish deposits on the tendons that occur in **familial hypercholesterolaemia**.

Xenobiotic A substance that does not originate in the body but is pharmacologically, endocrinologically or toxicologically active.

Xenogeneic transplant The use of a graft from another species. See also **Allograft**.

Zoonoses Diseases that can be transmitted from animals to humans.

Zymogens Precursor molecules that turn into active proteases when subjected to a specific proteolytic cleavage(s).

Taylor & Francis
Taylor & Francis Group
http://taylorandfrancis.com

Nature and investigation of diseases

1

OBJECTIVES

After studying this chapter you should be able to:

- define appropriate terms associated with health and disease
- give a simple classification of diseases
- list the ways pathological investigations help in the diagnosis, treatment and management of diseases
- discuss the clinical and analytical evaluations of clinical/pathology laboratory tests
- assess the role(s) of clinical laboratory audits

1.1 INTRODUCTION

Diseases and clinical conditions can be defined as any abnormality or failure of the body to function properly, which may require medical treatment (*Figure 1.1*). The scientific study of diseases is called **pathology**. Every disease has a distinct set of features that include its specific cause, for example pathogens and parasites (*Figure 1.2*), associated clinical symptoms and a characteristic progression, with related morphological and functional changes in the patient. The presence of an abnormality on its own, however, does not necessarily indicate disease, since the affected individual must also suffer from ill health.

Figure 1.1 An intensive care unit of a hospital. ⏎

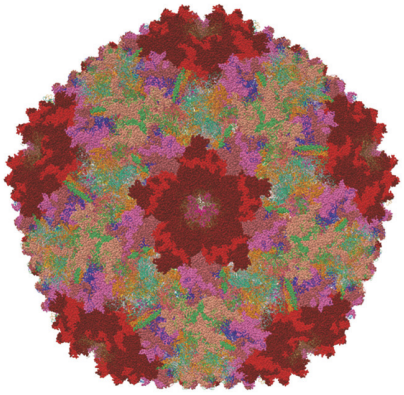


Figure 1.2 Viruses cause many infectious diseases. Molecular model of an adenovirus (diameter about 70 nm). This group of viruses cause respiratory infections. PDB file 6CGV. ⏎

Health can be defined as an absence of signs and symptoms associated with any disease. This definition has limitations in that there are circumstances where individuals believe they are ill even though detectable indications of disease are not present. Conversely, there are individuals who believe they are healthy but on detailed examination are found to suffer from a serious disease. For this reason, the World Health Organization (WHO) devised a more appropriate definition of health: 'a state of complete physical, mental and social well-being and not merely the absence of disease or infirmity'. Currently there is considerable interest, particularly in developed countries,

DOI: 10.1201/9781003296836-1

in promoting health by improving lifestyles and reducing the mental and social factors associated with ill health.

Individuals differ in their physical appearance and also in their composite biological materials. Differences between individuals are due to biological variation. Variation in values for measurable body features or biological substances (**analytes**) occur for a number of reasons. Differences in the genotypes of individuals ensure that no two individuals are the same (other than identical twins). Biological variation arises due to differences in lifestyle, experiences and diet, in addition to differences in DNA (*Figure 1.3*; genes are discussed in *Chapter 15*). There are also differences in physiological processes arising from bodily control mechanisms. For example, the concentration of glucose in the blood varies between individuals with diet, time of day and physical activity.

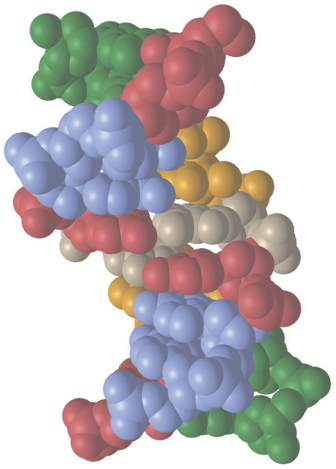


Figure 1.3 A molecular model of a decamer (10 base-pair length) of the genetic material, DNA. PDB file 1D61.

1.2 CHARACTERISTIC FEATURES OF DISEASES AND CLINICAL CONDITIONS

Each disease or clinical condition has a number of characteristic features, which allows them to be categorised and diagnosed and for appropriate treatment to be given to the patient. Such features include:

- aetiology
- pathogenesis
- clinical manifestations
- prognosis

Aetiology

Aetiology refers to the cause of a disease. Aetiological agents can be **endogenous**, in other words originating within the body, or **exogenous**, coming from outside the body. Endogenous agents include endocrine disorders and genetic defects (*Chapters 7* and *15*), while exogenous agents include prions, viruses and microorganisms, such as bacteria (*Figure 1.4*) and fungi that cause infections (*Chapters 2* and *3*), chemicals (*Chapter 12*), physical trauma and radiation (*Chapter 17*). Many diseases are said to be *predictable* and arise as a direct consequence of exposure to the causative agent. Others are considered *probable* in that they may be a consequence of the causative agent but the development of illness is not inevitable. An individual can be infected with a pathogenic microorganism but the outcome of the disease may depend on other factors such as the nutritional and immune status of the affected person.

Margin Note 1.1
Prions as endogenous and exogenous agents

Prions are aberrant native proteins that arise because of mutated genes and cause spongiform encephalopathies (*Chapters 2* and *15*); they are therefore endogenous agents of disease. However, prions can spread these diseases from patient to otherwise healthy individuals and are therefore infectious entities; thus, prions are also exogenous agents.

Some diseases have more than one aetiological agent and may, indeed, be caused by a range of factors. Such diseases are said to be **multifactorial** in origin. Diabetes mellitus type 2 (a disorder of carbohydrate, fat and protein metabolism) is believed to have a multifactorial origin involving several genetic, dietary and environmental factors (*Chapter 7*). Many diseases are of unknown cause and are said to be **idiopathic**. An example of this is hypertension, where more than 90% of cases are of unknown cause. The treatment of idiopathic diseases is restricted to alleviating the symptoms. Some conditions are caused by the effects of treatment and are called **iatrogenic diseases** (from the Greek word *iatros*, doctor). The treatment of some cancers with cytotoxic drugs, for example, can cause a severe iatrogenic anaemia, although they may be successfully treating the cancer.

Occasionally a disease of unknown aetiology is more commonly found in populations with certain dietary, occupational or lifestyle conditions called **risk factors**. Tobacco smoking is a significant risk factor in the development of heart disease and lung cancer (*Chapters 14* and *17*). Some risk factors may be important in the development of the disease, whereas others may make the individual more susceptible to disease.

Predisposing factors are conditions or situations that make an individual more susceptible to disease. They include age, sex, and heredity and environmental factors. For example, the immune system in a newborn is not fully developed and, as a consequence, babies are more susceptible to infections. However, during ageing, the immune system undergoes a progressive decline in function, making the elderly also more susceptible to infections. Sex may also be a predisposing factor: men are more likely to suffer from gout than women, whereas osteoporosis is more common in the latter.

Some diseases increase the risk of developing other clinical conditions. Thus, some conditions increase the risk of someone developing cancer and are said to be **premalignant**. This is seen in ulcerative colitis (*Chapter 11*), an inflammatory condition affecting the large intestine that increases the probability of developing bowel cancer in sufferers. Some diseases predispose the patient to other conditions by allowing infectious agents, not normally pathogenic, to cause disease. This is seen in **opportunistic infections** where a decline in the immunological functions of an individual makes them susceptible to

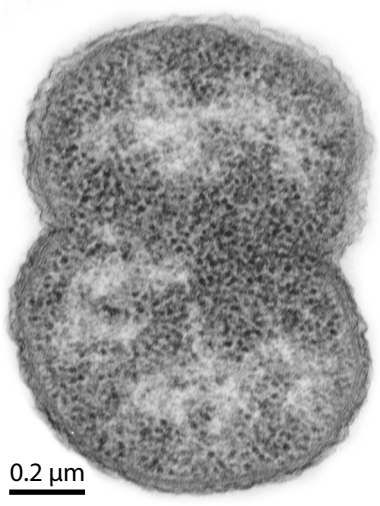


Figure 1.4 Electron micrograph of the bacterial pathogen *Neisseria meningitidis*, the causative organism of meningitis (*Chapters 2* and *3*). Courtesy of Dr A. Curry, Manchester Royal Infirmary, UK.

infections by microorganisms that are normally nonpathogenic. Human immunodeficiency viruses (HIVs; *Figure 1.5*) cause acquired immunodeficiency syndrome (AIDS); infected individuals (*Chapter 2*; see also *Box 3.3*) have little resistance to attacks by microorganisms, such as those responsible for causing pneumonia, as well as fungal organisms, such as the yeast *Candida*, which are part of the normal flora of the body (*Chapter 2*).

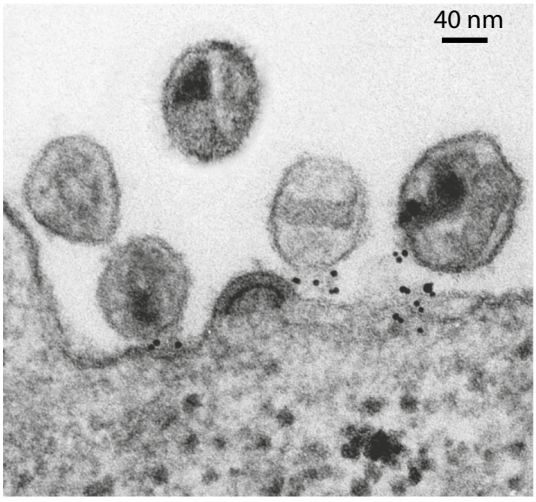


Figure 1.5 Electron micrograph of human immunodeficiency virus (HIV) particles escaping from a cultured human cell. Courtesy of Dr D. Robertson and Professor R.A. Weiss, Institute of Cancer Research, Royal Cancer Hospital, UK.

Diseases are often described as primary or secondary. Primary may refer to a disease of unknown cause (idiopathic), whereas secondary refers to conditions that arise from an existing disease. However, these terms are also often used to describe the stages of a disease. For example, in cancer, the primary tumour is the initial tumour; secondary tumours arise following its metastasis (*Figure 1.6*) to other tissues (*Chapter 17*).

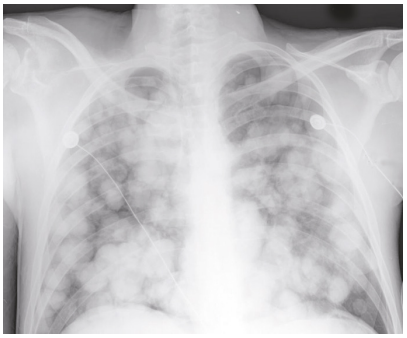


Figure 1.6 X-ray of a patient with numerous secondary tumours of the lungs. Courtesy of WikiDoc/Prashanthsaddala.

Pathogenesis

The **pathogenesis** of a disease describes its development or, more specifically, how the aetiological agent(s) act(s) to produce the clinical and pathological changes characteristic of that disease. Some examples of how diseases undergo pathogenesis include inflammatory reactions (*Chapters 4* and *5*) in response to harmful agents and carcinogenesis, where the formation of tumours occurs as a result of exposure to carcinogens (cancer-inducing substances).

Diseases have 'natural histories' that describe the typical patterns of how each disease usually progresses, its effects and its duration. The effects of the disease on the patient are referred to as **morbidity**. Occasionally, the morbidity of a disease may cause disability that, in turn, may restrict the activities of the patient. The **mortality** of a disease describes its possibility of causing death. This is usually expressed as a percentage. Some diseases have a rapid, often severe onset that is described as **acute**. However, other diseases are **chronic**, developing gradually over a relatively longer time. Thus, acute kidney disease is characterised by rapid deterioration of kidney function over a matter of days, while chronic kidney disease develops over months or even years (*Chapter 8*).

Diseases rarely occur immediately following exposure to an aetiological agent. In most cases, a period of time, the **incubation period**, must elapse before the disease becomes apparent. In carcinogenesis, this period may last several decades and is referred to as the **latent period**. With infectious diseases, the time between exposure and development of the disease is often characteristic of the infectious agent involved.

Clinical manifestations

Clinical manifestations of a disease are produced by its aetiology and pathogenesis: they include its characteristic **signs** and **symptoms**. Symptoms are indications that a disease is present and something of which the patient complains, for example nausea, malaise or pain (*Table 1.1*).

Signs are features that a clinician specifically looks or feels for, such as redness or swelling of the skin, when examining the patient. Some diseases present with a **subclinical** stage where signs and symptoms are not apparent, even though the disease is established. However, characteristic biochemical and cellular changes are detectable by laboratory analyses of, for example, blood or urine.

Symptom	Cause
Cough	Release of neuropeptides following irritation of the respiratory mucosa
Cyanosis	Reduced supply of oxygenated haemoglobin to the skin (see *Box 13.4*)
Diarrhoea	Inadequate absorption of food by the gastrointestinal system, which leads to an osmotic retention of water and production of watery stools (*Chapter 11*)
Fever	Interleukin-1 and prostaglandins stimulate the thermoregulatory centre in the brain
Pain	Stimulation of nerve endings by trauma, chemicals and/or heat
Swelling	Increase in the number or size of cells or an accumulation of tissue fluid
Weight loss	Decreased intake of food or a catabolic state stimulated by the release of factors from tumours

Table 1.1 The symptoms and causes of some diseases.

Clinical symptoms and signs are often accompanied by structural or functional abnormalities, called **lesions**, in affected tissues that are responsible for ill health and usually cause the symptoms and signs of disease. Lesions may be biochemical in nature, such as defective haemoglobin in patients with haemoglobinopathies (*Chapter 13*). Alternatively, a lesion may include the presence of abnormal substances in cells, tissues and organs, such as deposits (plaques) of amyloid in the brain in patients with Alzheimer disease (*Chapter 18*). Loss of healthy surface tissue, for example in gastric ulceration (see *Figure 11.40*), may also be a feature of some diseases. Diseases may involve functional abnormalities, such as an inappropriate secretion of hormones. Examples of these include the excessive production of cortisol in Cushing syndrome or the inadequate biosynthesis of growth hormone in dwarfism. Other functional defects might include impaired nerve conduction and muscular contraction.

The term **syndrome** is often used to describe certain diseases that are characterised by multiple abnormalities that form a distinct clinical picture. For example, Cushing syndrome occurs when an excess of the hormone cortisol produces a combination of clinical features and characteristic facial and body features including hirsutism (excessive growth of facial hair), hypertension and obesity. Some diseases frequently present with complications, that is, new or separate processes secondary to, and as a consequence of, the initial disease. Diabetic cataracts, retinopathy and nephropathy (*Figure 1.7*) can all be chronic complications of diabetes mellitus (*Chapter 7*).

The manifestations of a disease in different patients are not static and are affected by compensatory mechanisms in the body, as well as by environmental influences and, indeed, responses to treatment. Diseases often have a range of manifestations and their presence and severity may vary from patient to patient. In addition to variations between individuals, differences occur during stages of development, from infancy to old age. *Figure 1.8* outlines some key features of diseases.

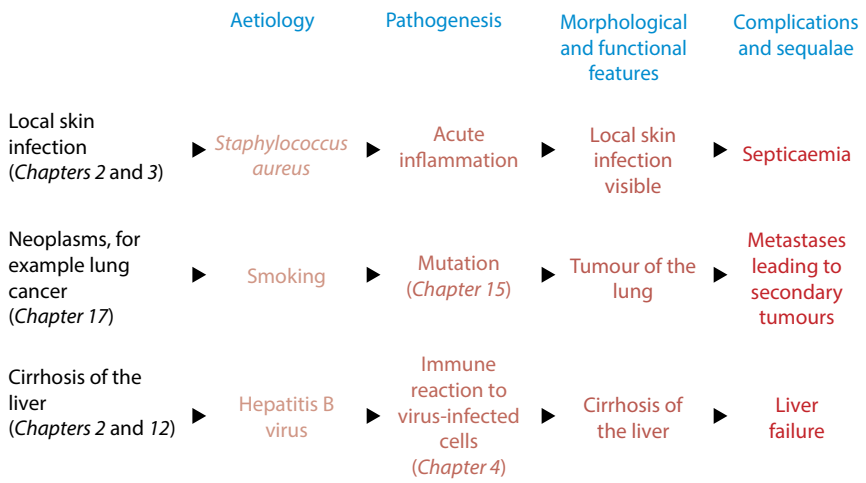


Figure 1.8 Characteristics of some diseases showing the relationships between aetiology, pathogenesis, morphological and functional features, and complications and sequelae.

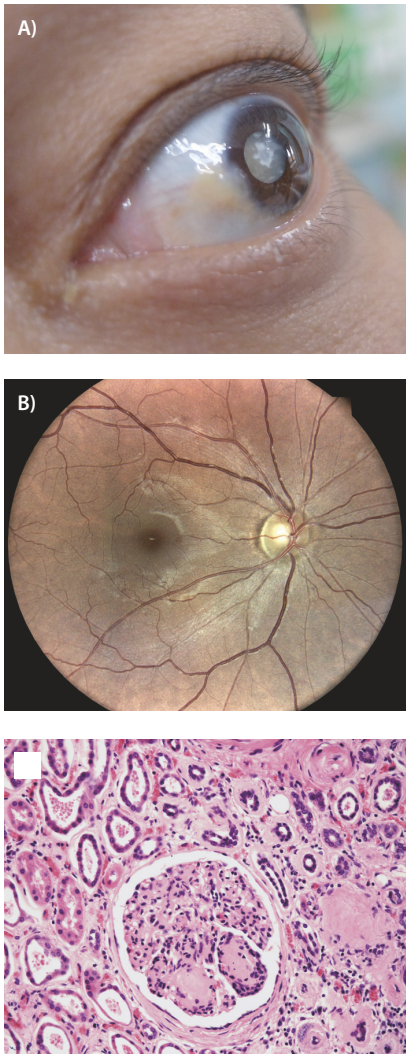


Figure 1.7 Some of the chronic complications of diabetes mellitus. (A) Mature cataract secondary to diabetes mellitus affecting a young adult. (B) Background diabetic retinopathy on right eye. (C) Advanced diabetic nephropathy: the glomerulus shows diffuse mesangial expansion, which occurs because of excessive deposition of extracellular matrix proteins such as fibronectin in the mesangium, a cellular membrane that supports the capillaries of the glomerulus. Intercapillary glomerulosclerosis is visible, and the Bowman capsule basement membrane is thickened. (A), Courtesy of Shutterstock.com/Casa nayafana; (B), courtesy of Shutterstock.com/DrIsfan; (C), courtesy of Shutterstock.com/Jose Luis Calvo.

Prognosis

Prognosis is the likely outcome of a disease. Prognoses can vary considerably between different diseases and, of course, can be influenced by treatment. Hence, when giving a prognosis, it is necessary to clarify if the disease is following its natural course or whether there is, or needs to be, medical or surgical intervention. A viral disease such as German measles (rubella) will normally resolve of its own accord, whereas a broken leg or a heart attack needs treatment. With some diseases, especially some cancers, patients may go through a period of good health with a reduction or disappearance of the symptoms and the disease is said to be in **remission**. However, a **relapse** may occur, with a return of the disease symptoms following this period of apparent recovery. Diseases with tendencies towards remission and relapse include ulcerative colitis and acute lymphoblastic leukaemia (*Chapters 11* and *17*).

1.3 CLASSIFICATION OF DISEASE

Some diseases share common features and can be grouped together in a classification system. One way of classifying diseases is on the bases of their causes (*Table 1.2*). However, this system is not perfect, as some diseases have multiple causes and there is likely to be an overlap between the different categories.

Congenital diseases are present at birth and may or may not have been inherited. The constitutive developmental defect may be known or of an unknown cause, and although they may be genetic in origin (see below), others could be environmental. For example, fetal alcohol syndrome is a congenital condition arising from an excessive alcohol intake by the mother during pregnancy.

Degenerative diseases involve the progressive loss of body tissues and impairment of their functions. They are often associated with mutations or ageing. Examples include the relatively common muscular dystrophies and neurodegenerative diseases, such as Alzheimer and Parkinson diseases (*Chapters 16* and *18*).

Endocrine diseases arise from the under- or overproduction of hormones or from resistance to a particular hormone, perhaps because its cellular receptor is absent as the result of a mutation (*Chapter 7*). Thus, for example, acromegaly is caused by the overproduction of growth hormone (see *Figure 7.25*) in adults, while type 2 diabetes mellitus is a consequence of insulin resistance, when appropriate target cells fail to respond to the hormone (*Figure 1.9*).

Genetic diseases are due to defects in the genes or chromosomes of individuals (*Chapter 15*). A defective gene may result in the inadequate production of a key enzyme, such as phenylalanine hydroxylase in phenylketonuria (see *Figure 15.20*). Down syndrome is an example of a disorder caused by an abnormal chromosome complement (*Chapter 15*). Some genetic disorders are not inherited from parents and may arise from a new genetic mutation in the offspring, as in the disease progeria (see *Box 18.3*).

Homeostatic diseases arise when mechanisms for controlling homeostasis are disrupted. For example, in inappropriate antidiuretic hormone (ADH) secretion (*Chapter 8*), a diminished urine production leads to an increase in body fluids.

Iatrogenic diseases occur as a consequence of clinical treatment. For example, patients who are receiving drugs, such as thiazide diuretics (see *Figure 9.11*) to control their blood pressure, may suffer from a low concentration of K^+ in the serum (hypokalaemia; see *Chapter 8*) caused by excessive renal losses. If untreated, hypokalaemia may, in turn, cause cardiac arrhythmias (*Chapter 14*).

Idiopathic diseases are those of unknown cause.

Immunological diseases occur in circumstances in which the immune system can damage the body's own tissues. In autoimmune conditions, for example autoimmune thyroiditis, antibodies are produced that attack the patient's own tissues. Alternatively, there are diseases associated with immunodeficiency that increase the susceptibility of the patient to infectious agents. This occurs in severe combined immunodeficiency (SCID; see *Chapter 5*) and in HIV–AIDS.

Infectious diseases are caused when prions, viruses, bacteria, protozoa, fungi and helminths enter and spread within the body.

Neoplastic diseases are characterised by the uncontrolled and abnormal growth of cells. These cells form benign or malignant tumours. Malignant neoplasms are a major cause of death in many developed countries. Moreover, their incidence is increasing as people live longer.

Nutritional diseases result from an inappropriate intake of nutrients, such as proteins (which supply essential amino acids), carbohydrates, essential fatty acids, vitamins or trace elements. Inadequate nutrition is a major cause of disease, particularly in developing countries. Such deficiencies may be generalised, as in protein–energy malnutrition

Class of disease
Congenital
Degenerative (*Chapter 18*)
Endocrine (*Chapter 7*)
Genetic (*Chapter 15*)
Homeostatic
Iatrogenic
Idiopathic
Immunological (*Chapter 4*)
Infectious (*Chapters 2* and *3*)
Neoplastic (*Chapter 17*)
Nutritional (*Chapters 10* and *11*)
Psychogenic
Toxic (*Chapter 12*)
Traumatic

Table 1.2 A classification of diseases.

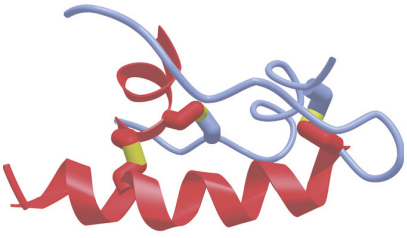


Figure 1.9 Model of insulin. PDB file 1ZEH. See also *Figures 7.3* and *7.40*.

where there is simply not enough food, or there may be a lack of a specific nutrient, for example vitamin A (*Figure 1.10*), leading to several disorders, including night blindness. In contrast, in many developed countries, an excessive intake of energy combined with a lack of exercise is responsible for a worrying increase in obesity (*Chapter 10*).

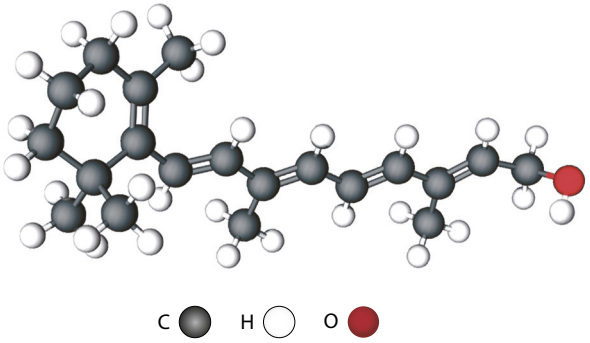


Figure 1.10 Computer drawn model of vitamin A (retinol). See also *Figure 10.19*.

Psychogenic diseases originate in the mind. They may have a significant psychological or emotional component as seen, for example, in schizophrenia.

Toxic diseases are caused by the ingestion, inhalation or absorption through the skin of a variety of poisons that may be encountered in the environment. For example, carbon monoxide can be inhaled from car exhausts or faulty gas fires or water heaters and cause tissue hypoxia and death (*Chapter 12*). Exposure to toxic agents may be accidental or deliberate.

Traumatic diseases are caused by physical injury and include mechanical trauma, extremes of heat or cold, electric shocks and radiation. Apart from the obvious problems caused by extensive damage to tissues, traumatic diseases may render an individual more prone to infection by compromising the functions of the immune system.

1.4 EPIDEMIOLOGY OF DISEASES

Epidemiology is the study of how diseases spread in populations in relation to their causal factors. Consequently, it is largely concerned with the collection and interpretation of data about diseases in groups of people rather than in individuals. The types of data collected in epidemiological studies provide information about, for example, the **prevalence** and **incidence** of a disease. Prevalence refers to the proportion of people in a population affected at a specific time. The incidence is the number of new cases of a disease in a population occurring within a specified period.

Data about the aetiology of diseases indicate if there is a need for screening or the introduction of other preventative measures, and whether the prevailing healthcare facilities are appropriate. Also, epidemiological studies can often provide information about the cause(s) of diseases. Thus, if a disease has a high incidence in a particular region or discrete population, then the disease may be caused by environmental factors peculiar to that area or it may have a genetic origin. Epidemiological studies of migrant populations are especially useful, since they can provide valuable information on the aetiology of a disease. A case in point might be a migrant population with a high incidence of a particular disease that then moves to another geographical area where the incidence of the same disease is low. If the incidence of disease in the migrant population remains high, then it is likely that the disease has a genetic basis. If, however, the incidence in the migrant population decreases to the level of the new geographical region, then environmental factors at the original site probably played a role in the aetiology.

The data on the incidence of some diseases are reliable. This is especially so for some infectious diseases and cancers that are **notifiable**. Clinicians are legally required to supply details of all new cases of diseases on the notifiable list to a central register. However, obtaining data on the incidence of other diseases can be difficult. For most diseases, the data obtained refer to mortality rates for that disease based on the causes listed on death certificates. This method of obtaining data has the major limitation of underestimating the incidence if the disease does not have a fatal outcome or overestimating it when several diseases are listed as contributory factors on death certificates.

The incidence of certain diseases changes with time and also can vary considerably from one country to another and even within different regions of the same country.

These differences are particularly marked between developing and developed countries. Infectious diseases and malnutrition are still more prevalent in developing countries, while, in the developed world, the incidence of many infectious diseases has been reduced dramatically in the last 100 years. The infant mortality rate is often used as a measure of health related to socioeconomic status. In general, the infant mortality rate is larger in developing compared with developed countries.

The decreased incidence of many diseases in developed countries may reflect changes in exposure to causative agents and the effects of preventative measures. For example, the reduction of diseases such as cholera is associated, in a large part, with improved public health measures. Improvements in sanitation, sewage and hygiene have had a considerable impact in reducing the incidence of many infectious diseases (see *Figure 2.2*). Mass immunisation against infectious diseases, such as polio, has had enormous beneficial effects in reducing disease in a population (see *Box 3.4*). Unfortunately, the reduction in infectious diseases has been accompanied by an increasing incidence of other diseases, such as cardiovascular diseases, diabetes, several types of cancers and psychiatric diseases. All are associated with ageing and, to a certain extent, this may reflect the increased life expectancy in developed countries: people are not killed by infectious diseases and live longer. Some evidence does suggest that the increased incidence of these diseases is also due to changes in diet, such as increased consumption of saturated fats, and other lifestyle factors, for instance a lack of exercise. **Intervention studies** aimed at changing diet and lifestyle factors in an attempt to reduce the incidence of these diseases are already proving beneficial.

Socioeconomic factors also influence the incidence of many diseases. Poverty tends to be associated with an increased incidence of malnutrition, and malnourished individuals are more susceptible to infectious diseases. Overcrowding is known to promote the spread of infectious diseases resulting in epidemics.

Some diseases have a high incidence in populations associated with specific occupations. For example, coal miners have a high incidence of pneumoconiosis caused by inhalation of coal dust and, in the past, those working with asbestos faced a high risk of asbestosis and of developing mesothelioma of the lung. Occupational hazards need to be identified and minimised to reduce the incidence of these diseases.

1.5 INVESTIGATING DISEASES

For most diseases, the clinical outcome, that is, prognosis, will be improved if treatment is started at an early stage. Consequently, an appropriate clinical investigation of disease is necessary to ensure a rapid and accurate diagnosis and to allow appropriate treatment to be initiated as soon as possible.

The procedure for investigating a disease is outlined in *Figure 1.11*. It starts with the affected person presenting symptoms and visiting his or her physician when feeling unwell or after a period of ill health. The examination usually begins with the clinician asking the patient about his or her current and past medical histories; current and previous medications; the use of alcohol, tobacco and/or other drugs; any family history of disease; and possibly their occupational history. This is usually followed by a clinical examination to look for signs of any lesion. This may involve visual examinations of the skin, eyes, tongue, throat, nails and hair, together with tests to assess cardiovascular, respiratory, gastrointestinal, genitourinary, nervous and musculoskeletal functions.

Since diseases typically present with recognisable symptoms and signs, the clinician may make a diagnosis of the disease based on the clinical history and the examination and then initiate treatment. However, this may not be possible, given that many clinical symptoms and signs are not specific to any one disease. The clinician may only be able to make a provisional diagnosis or shortlist several possible diagnoses. However, a range of diagnostic investigations is available to the clinician at the surgery, clinic or hospital in modern well-funded healthcare systems. Clinical/pathology laboratory investigations and physiological function tests can be requested and the results used to confirm, reject or distinguish between the various provisional diagnoses (*Section 1.7*). Increasingly, imaging techniques are used to investigate diseases and clinical conditions. X-rays, of course, have had a long history of clinical use since their discovery (see, for example, *Figure 1.6*), while an ultrasound scan, sometimes called a sonogram, is widely used (see, for example, *Figure 14.11*) but is perhaps best known to the wider public for the intrauterine monitoring of pregnancies (*Figure 1.12*). However, as their costs have reduced, computed tomography (CT) scanning, magnetic resonance imaging (MRI) and positron emission tomography (PET) have become significant radiographic tools (*Figure 1.13*; see also *Figures 14.12*, *18.18* and *18.19*).

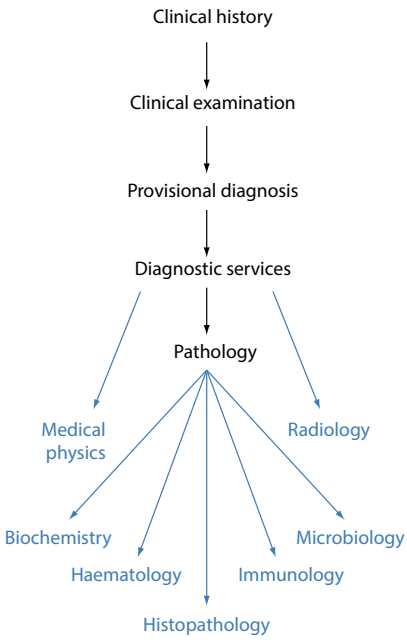


Figure 1.11 An outline of the steps involved in investigating diseases.

Figure 1.12 An ultrasound scan image of the head of a 30-week old fetus. Courtesy of the mother.

Margin Note 1.2
Röntgen and the discovery of X-rays

Wilhelm Röntgen (1845–1923) was a German engineer and physicist who first detected X-rays (sometimes called Röntgen rays) in 1895. Following their discovery, X-rays were rapidly introduced into diagnostic medicine. Röntgen received the inaugural Nobel Prize in Physics in 1901 for his discovery.

1.6 TYPES OF CLINICAL/PATHOLOGY LABORATORIES

The functions of hospital clinical/pathology laboratories are to scientifically investigate diseases to aid their diagnoses, for screening purposes, and to monitor and assess the progress of any treatment of the disease. The typical pathological laboratory service disciplines offered by hospitals are:

- clinical biochemistry
- clinical genetics
- haematology
- histopathology
- immunology
- medical microbiology

Clinical biochemistry is concerned with investigating the biochemical changes associated with diseases. A wide range of analytes are measured in clinical biochemistry laboratories. Some of these analyses are carried out routinely on all samples (blood, urine) coming into the laboratory using automated methods (*Figure 1.14*); others need to be requested especially.

Figure 1.14 An automated analyser in a hospital pathology laboratory capable of performing most of the major clinical investigations. Courtesy of L. Jones and Drs C. Chaloner and E. Cowdell, Manchester University NHS Foundation Trust, UK.

Analyses include those for proteins, enzymes, hormones, lipids, tumour markers, blood gases, sugars and inorganic ions to investigate a variety of disorders, including those associated with abnormal renal, respiratory, metabolic, bone and endocrine function. In addition, analytes are measured during investigations of genetic disorders, both to diagnose and to monitor the effectiveness of therapies.

Clinical genetics is a growing area in the investigation of diseases. A major focus of clinical genetics laboratories is the identification of genetic abnormalities (*Chapter 15*). This could include, for example, identifying the number and form of chromosomes (*Figure 1.15*) in blood films to identify any numerical and structural abnormalities. Techniques, such as fluorescence *in situ* hybridisation (FISH; see *Margin Note 5.4*) can detect small specific chromosomal lesions that would otherwise not be observable, and so can identify genetic defects associated with some cancers (*Figure 1.16*).

Haematology is concerned with the study of disorders of blood, including blood-clotting (coagulation) defects (*Chapter 13*). Haematological investigations can involve determining the concentrations of blood proteins, such as haemoglobin, to aid in the diagnosis of diseases. The microscopic examination of blood and marrow films, that is, thin layers of these tissues, spread out on a microscope slide and stained (*Figure 1.17*) may also be helpful.

Some haematology laboratories may also be involved in the provision of blood and blood products for transfusion services, but these are often run as separate services (see *Chapter 6*).

Histopathology is concerned with investigating diseases by examining cells and tissues from patients. This involves the macro- and microscopic investigation of body

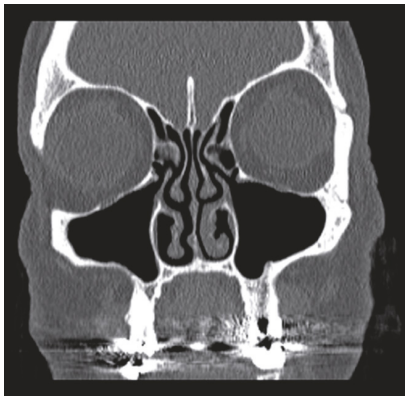


Figure 1.13 A computed tomography (CT) scan of the front of the skull taken when investigating a possible occlusion of the paranasal sinuses and which confirmed an absence of sinonasal disease. Courtesy of M. Khalid-Raja, Consultant Otolaryngologist, Stepping Hill Hospital, UK.

Figure 1.15 A spread of human chromosomes from a female.

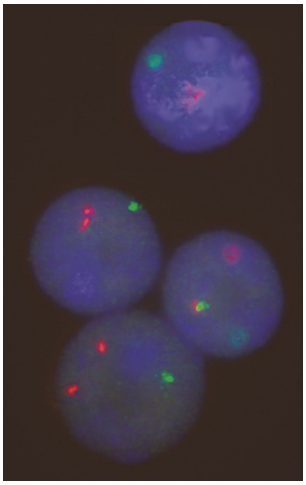


Figure 1.16 The use of fluorescence *in situ* hybridisation (FISH) to diagnose chronic lymphocytic leukaemia. The test uses two probes to detect deletion mutations (*Chapter 15*). One probe fluoresces green and binds to mutated ataxia-telangiectasia (*ATM*), a gene that normally encodes a serine/threonine kinase that is recruited to sites of DNA double-stranded breaks. The other, red-fluorescing probe detects deletion mutations in the gene for the transcription factor p53, which responds to diverse cellular stresses to regulate target genes that induce cell cycle arrest and apoptosis (see *Chapter 17*). A wild-type cell pattern would be two green and two red signals. However, in this bone marrow aspirate sample, a cell with one green and two red probe signals is visible, unfortunately indicating an AT deletion. Courtesy of Dr Nick Telford and Clare Hodgson, Oncology Cytogenetics, The Christie Pathology Partnership, The Christie NHS Foundation Trust, Manchester, UK. ⏎

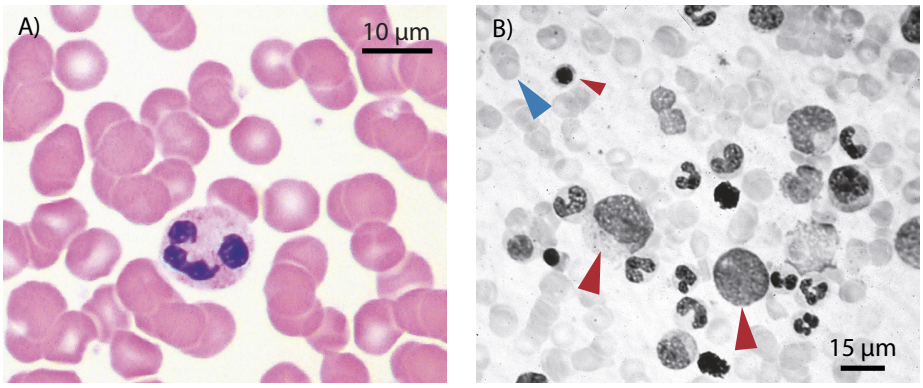


Figure 1.17 Photomicrographs showing (A) a blood film with a single leukocyte surrounded by erythrocytes (*Chapter 13*) and (B) bone marrow, that is, the site of erythrocyte and leukocyte formation, showing its normal range of immature erythrocytes (smaller red arrowheads) and leukocytes (blue arrowheads). The large proportion of mature erythrocytes (larger red arrowheads) in the background is due to unavoidable contamination with peripheral blood during sample collection. Courtesy of Dr L. Seal and J. Overfield, School of Biology, Chemistry and Health Science, Manchester Metropolitan University, UK. ⏎

tissues for the identification of diseases. Among other things, histopathology laboratories are usually involved in the diagnosis of malignancies but they can also provide information on how far a tumour has progressed ('staging') and therefore can suggest a likely prognosis. In addition, histopathology laboratories may also assist with investigation of a range of infectious and inflammatory conditions affecting body tissues.

Immunology laboratories are concerned with studying the immune system's response in both healthy and diseased states (*Chapters 4* and *5*). Immune responses are 'cell-mediated' or 'humoral'. The former involves T lymphocytes and the latter the production of antibodies by specialised B lymphocytes. The presence or absence of antibodies in plasma can be determined, for example, by serum electrophoresis to assess generalised immunodeficiencies and other diseases (*Figure 1.18*). However, of more diagnostic value during the investigation of immune diseases may be the measurement of specific antibodies that are produced in response to a particular antigen (which may be an infectious agent or an autoantigen). The numbers of cells involved in immunity, such as T and B cells and T helper and T suppressor cells (*Chapters 4* and *5*), are often determined, as this can provide valuable information about the immune status of an individual.

Medical microbiology is concerned primarily with the detection and identification of clinically pathogenic viruses, bacteria, protozoa, fungi and helminths (worms). Viruses can be observed using electron microscopy techniques and therefore identified in clinical specimens. However, if larger titres of viruses are required, they must be cultured in eukaryotic cells (*Chapter 2*). Microorganisms are detected directly in specimens obtained from the patient on simple swabs (see *Figure 1.25*) that are cultured in growth medium to increase the number of microorganisms and allow their easier detection. The

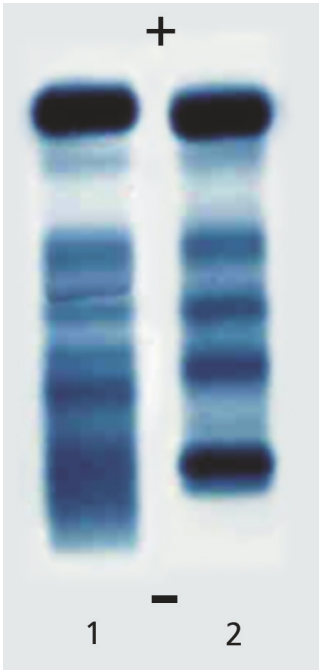


Figure 1.18 Excessive immunoglobulin (*Chapters 4* and *5*) is produced in multiple myeloma, a tumour of B lymphocytes, which produce immunoglobulins. The proteins in samples of serum can be separated by electrophoresis and visualised by staining them with a dye. Lane 1 shows the separated proteins from 'normal' serum. Serum albumin is the most abundant protein and is the strongly staining band near the positive end. Lane 2 shows a myeloma serum sample with a second dense band at the negative end. This band shows the relatively enormous amount of a single type of immunoglobulin produced by the tumour. In multiple myeloma, the serum protein concentration increases, making the blood viscous and straining the heart and kidneys. In addition, so much effort is put into the biosynthesis of one type of useless antibody that the concentrations of other antibody molecules decreases and patients become prone to infections. ⏎

presence of microorganisms may also be determined indirectly by detecting antibodies produced by the patient in response to the infection. Medical microbiology laboratories also investigate the responses of pathogenic microorganisms to antibiotics.

1.7 HOSPITAL LABORATORY TESTING

Tests performed by the clinical/pathology laboratories can assist clinicians in investigating diseases. Some tests may only give a subjective assessment, for example when a pathologist assesses the types of cells obtained from a fine needle aspirate of a suspected breast tumour. However, others can provide quantitative information, such as the concentration of thyroid hormones in the serum, which can then be compared with a **normal** value. Unfortunately, the term 'normal' is often difficult to define in clinical terms. To alleviate this problem, **reference ranges** have been widely adopted (*Box 1.1*). Numerical reference limits are based on the mean value plus or minus two standard deviations (SDs) against which test results can be compared. The term **normal range** is often used synonymously with reference range.

BOX 1.1 REFERENCE RANGES

The reference range for any particular analyte can be obtained by measuring it in healthy individuals from a representative sample of the local population. Most laboratories use healthy blood donors. It may be necessary to know the normal range in adults, children, males, females, particular ethnic groups, and pregnant or postmenopausal women. However, since no analytical method is 100% accurate and individuals vary, care must be taken when establishing a reference range. To determine the reference range, the values for the measured analyte are plotted against the frequency of the analyte in the selected population. In most cases, the resultant graph shows a normal or Gaussian distribution, with most values clustered around its centre, and from which the mean ($\bar{x}$) and standard deviation (SD) can be determined (*Figure 1.19*). The $\bar{x}$ is the arithmetic average value of a particular group of measurements and can be calculated using the formula:

$$\bar{x} = \sum x / n$$

where n is the number of individual measurements and $\sum$ is the total of individual values x. The SD is a measure of the dispersion of these data. It is defined as the square root of $\sum (x - \bar{x})^2 / n - 1$. Hence, the smaller the value of the SD relative to that of the mean, the less dispersed the data. It can be seen from *Figure 1.19* that the $\bar{x} \pm$ SD includes about 68% of the samples; the $\bar{x} \pm 2$SD will include approximately 95% of the samples. The coefficient of variance (CV) is also a measure of dispersion. It is related to both the $\bar{x}$ and SD:

$$CV = 100 \times SD / \bar{x}$$

Since both the $\bar{x}$ and the SD have the same units, the CV is a percentage and the lower its value, the smaller the dispersion.

In general, a reference range is taken as the value two SDs either side of the $\bar{x}$, which, provided the curve is Gaussian, will cover 95% of the values obtained for the selected sample. The 95% reference range was selected, as this minimises any overlap between the results for a healthy population and those for a population with the disease. However, choosing a 95% range is one of the major limitations of reference ranges, since 5% of healthy individuals will, by definition, give results that are outside these values. Thus, a test result outside the reference range does not necessarily imply that the individual is ill, although it does indicate that there is a greater likelihood of the presence of disease.

The profile of test results for some substances, for example serum bilirubin (*Figure 1.20*), does not give a Gaussian curve but shows a skewed distribution. However, this skewed

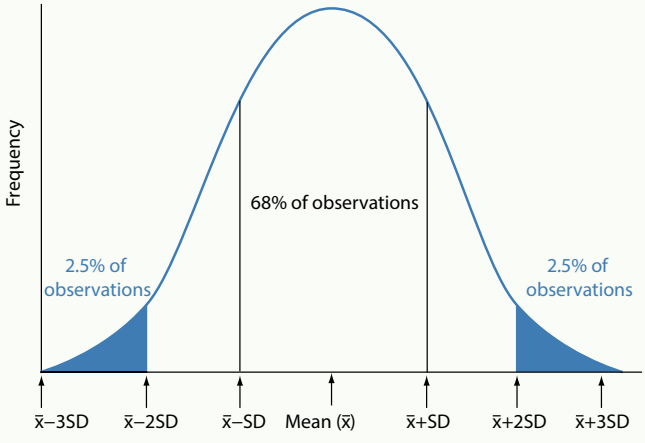


Figure 1.19 Gaussian distribution of results from a test on a healthy population. The reference range encompasses 95% of these results, which are found within minus and plus two standard deviations of the mean ($\bar{x}$–2SDs and $\bar{x}$+2SDs).

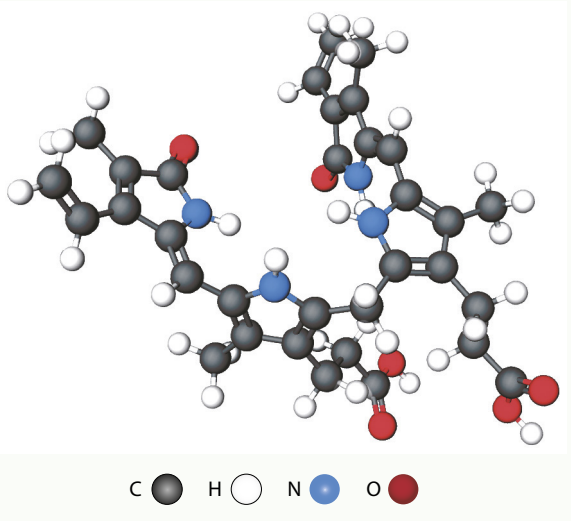


Figure 1.20 Computer drawn model of bilirubin. See also *Figure 12.5*.

distribution can be transformed mathematically to a Gaussian distribution and a normalised reference range calculated.

The values of several analytes, such as serum iron and alkaline phosphatase, vary with the age or sex of the patient. In such cases, age- and sex-matched reference ranges are required. When interpreting results for a particular patient, the ideal reference range would be obtained from the same patient before their illness and this is sometimes possible. For example, the concentrations of electrolytes in serum can be measured in a patient before an operation for comparison with those obtained postoperatively. However, in most cases, the results for the patient before they became ill are not available.

Roles of hospital clinical laboratory testing

In general, the roles of laboratory tests are to:

- identify the disease, that is, assist diagnosis
- monitor treatment
- be used for screening and the assessment of risk
- inform the physician and the patient of the likely outcome, that is, assist prognosis
- detect disease complications

The results from laboratory investigations are used in combination with the clinical history and examination of the patient to determine the nature of the disease. Thus, a low value for the concentration of glucose in the blood of a patient can confirm hypoglycaemia and the clinician can start palliative treatment, even though the cause of the hypoglycaemia may be unknown at this stage.

Laboratory tests may be used to monitor the course of an illness or the effects of its treatment. For example, the concentration of glycated haemoglobin (HbA_{1C}) in erythrocytes is measured in diabetics on a regular basis. The higher the concentration of glucose in the serum, the more readily the sugar reacts with proteins in a nonenzymic reaction to form HbA_{1C}, in which sugar molecules are covalently attached to the protein (see *Figure 18.7*). Hence, the amount of HbA_{1C} is an indicator of average glycaemia in such patients over the last three to four months. Diabetics who are not complying with their treatment by not taking their insulin regularly or giving themselves the wrong dose, or whose treatment is ineffective, can be identified because poor control of blood glucose gives rise to higher concentrations of HbA_{1C} than is normal or even found in well-controlled diabetics (*Figure 1.21*).

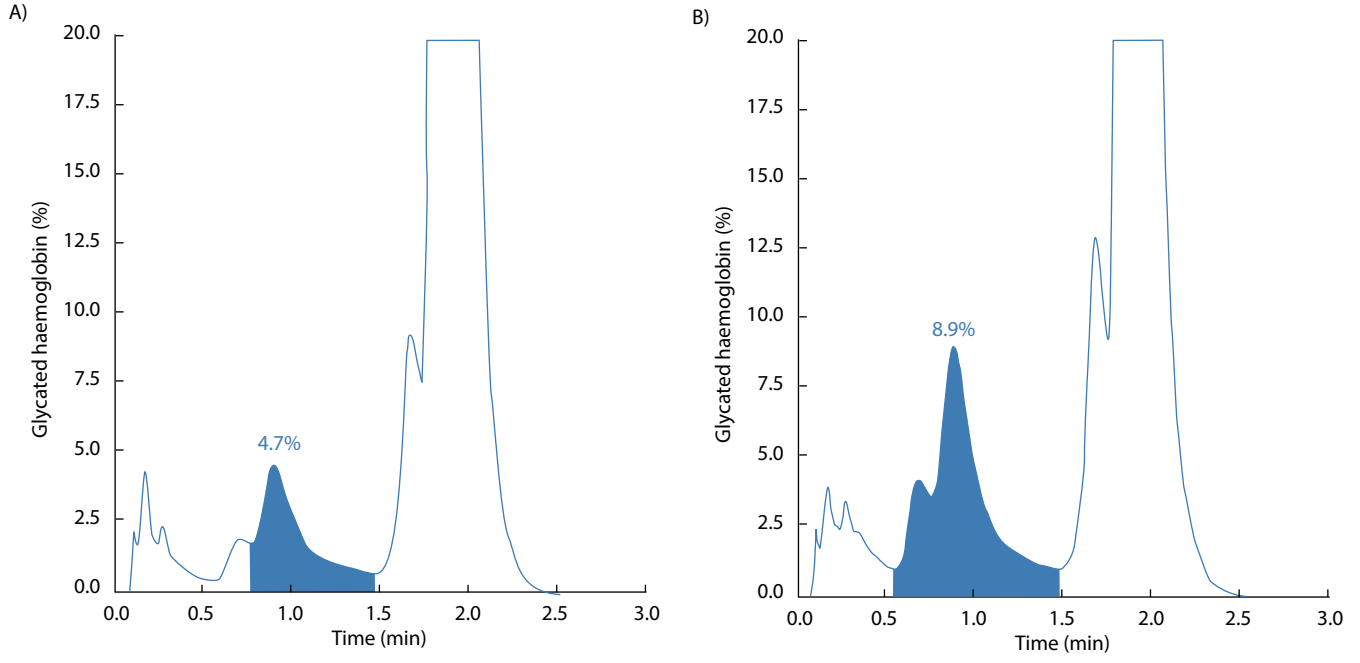


Figure 1.21 Chromatograms of glycated haemoglobins (HbA_{1C}) determined by high-performance liquid chromatography (HPLC). The results are shown for (A) a healthy person and (B) a patient with diabetes mellitus. Courtesy of the Department of Clinical Biochemistry, Manchester Royal Infirmary, UK.

Laboratory tests may be able to detect a disease *before* it presents clinically. This is referred to as **screening**. For example, the concentration of phenylalanine in the serum of all newborn babies in the UK is measured to detect phenylketonuria (*Chapter 15*). Affected children have high concentrations of serum phenylalanine (hyperphenylalaninaemia),

and metabolites of phenylalanine, such as phenylpyruvic acid, are also usually present (see *Figure 15.20*). If untreated, this condition leads to irreversible brain damage but, if caught early and treated with an appropriate diet, individuals develop normally. Other examples of screening tests include those for cervical and bowel cancers (*Margin Note 1.3*). Screening tests may also be used in certain groups of people to assess occupational exposure to harmful substances such as lead and radiation.

Laboratory tests can be used to assess the risk of developing a disease. The risk of developing coronary artery disease rises with increasing concentrations of blood cholesterol, and more so if other risk factors, such as smoking, obesity or diabetes, are present.

Tests in pathology laboratories can aid prognoses by indicating the likely outcome of a disease. Renal failure (*Chapter 8*) is a progressive disease that leads to a gradual build-up of creatinine in the serum (*Figure 1.23*). Measurements of serum creatinine may therefore indicate end-stage renal disease when the patient may need dialysis to survive.

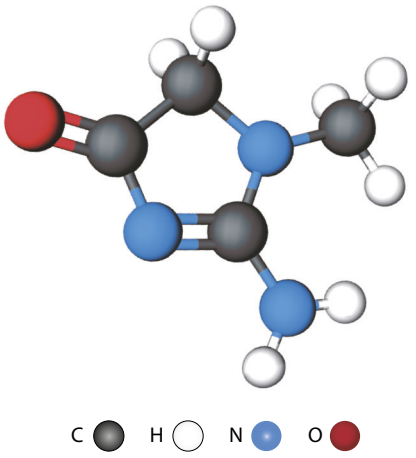


Figure 1.23 Computer drawn model of creatinine, which is formed from creatine phosphate by the body at a relatively constant rate and excreted in urine. It is produced in amounts that are essentially proportional to muscle mass and so its concentration in blood is commonly used as an indicator of kidney function (*Chapter 8*).

Vital information regarding the development and complications in a particular disease may be provided by laboratory tests. Urine is normally essentially free of protein; hence, the presence of 30–200 mg/L of serum albumin in the urine (microalbuminuria) of diabetic patients may indicate the development of nephropathy, a common secondary complication of diabetes.

Hospital laboratory tests

The wide range of clinical tests routinely offered by hospital laboratories must undergo a thorough evaluation for both analytical and clinical performances. Also, all laboratory staff must be appropriately trained to ensure high analytical standards and that they are able to produce valid test data.

Analytical methods need to satisfy certain criteria to ensure they are practical and suitable for use in the laboratory. For example, they must be accurate, precise, safe to use and of low cost. In addition to thorough evaluation of analytical methods, other aspects, such as the stability of samples, need to be considered. For example, some samples must be assayed immediately, while others can be stored at an appropriate temperature.

Analysers (see *Figure 1.14*) require checking at regular intervals to ensure analytical methods are working correctly and reproducibly. This usually means analysing control aliquots of urine and serum. Many of the companies that sell clinical apparatus and reagents also supply standard sera for example, which are tested in laboratories across the country. The results are recorded in nationwide databases so that comparisons can be made easily and laboratories can check that their methods are all giving similar or acceptable results.

Hospital laboratories have witnessed and are continuing to witness increasing workloads in processing large numbers of specimens daily, although they are helped by automation (see *Figure 1.14*). However, like all organisations, clinical establishments have finite budgets, and it is essential that the cost of tests is kept to a minimum. As well as direct expenses, such as those for reagents, equipment and labour, there are indirect costs, for example the heating and lighting of laboratories. Safety is also of importance and biological, chemical, mechanical and electrical hazards associated with the method need to be assessed to ensure the safety of hospital laboratory staff. New methods are introduced into the laboratory only if they offer significant advantages over existing methods. A clinical demand for the test must be established, and its clinical relevance is subject to review. New methods are assessed for their speed, that is, how many specimens

Margin Note 1.3
Cervical and bowel cancer screening tests

Cervical smear (*Figure 1.22*) testing is a screening that looks for abnormal changes in cells of the cervix, that is, the neck of the uterus (*Chapter 17*). Some of these abnormal cells can develop into cancer over 10 or more years. The most common cervical cancer, squamous cell carcinoma, is largely preventable given that treatment of the abnormal cells will remove them in more than 90% of cases, although, occasionally, further treatment may be needed. The development of a precancerous state in the cervix is described more fully in *Chapter 17*.

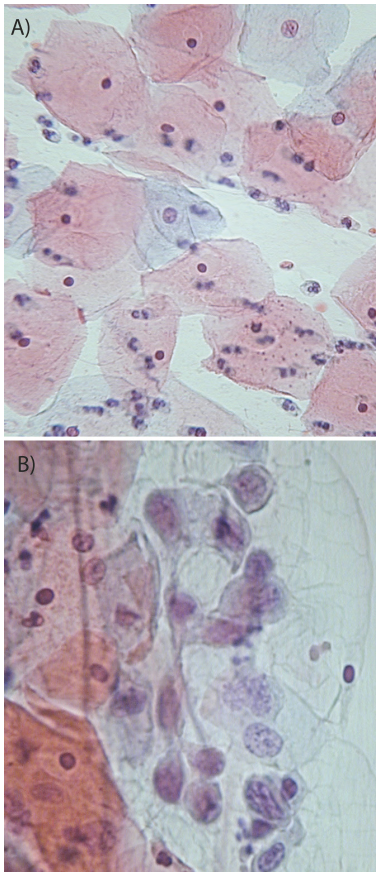


Figure 1.22 (A) Photomicrograph showing normal squamous cells in a cervical smear from the superficial layer of the cervix and from the layer of the cervical wall immediately below that of the squamous cells. The cells are healthy and have comparatively small nuclei. (B) Photomicrograph showing abnormal cells from a different patient. Note the comparatively large nuclei compared with the healthy cells. Courtesy of H. Glencross, Manchester Cytology Centre, Manchester Royal Infirmary, UK.

In the UK, bowel cancer screening is conducted in people over the age of 50 years using a faecal immunochemical

can be processed in a given time and how long it takes to produce a result. The time a test takes may, of course, be vitally important in the care and treatment of a patient when urgent intervention is necessary.

Requesting a clinical test

Figure 1.24 illustrates the overall procedure routinely followed when a clinician requires a hospital test. Note that this procedure can be divided into a number of distinct steps:

- a request for the test is made using a paper form or email, with patient details recorded
- the specimen is collected, labelled and transported and storage instructions are given
- the analysis is undertaken, that is, test results are obtained
- the results are interpreted, which are often printed out with the reference range values to help the clinician make an interpretation

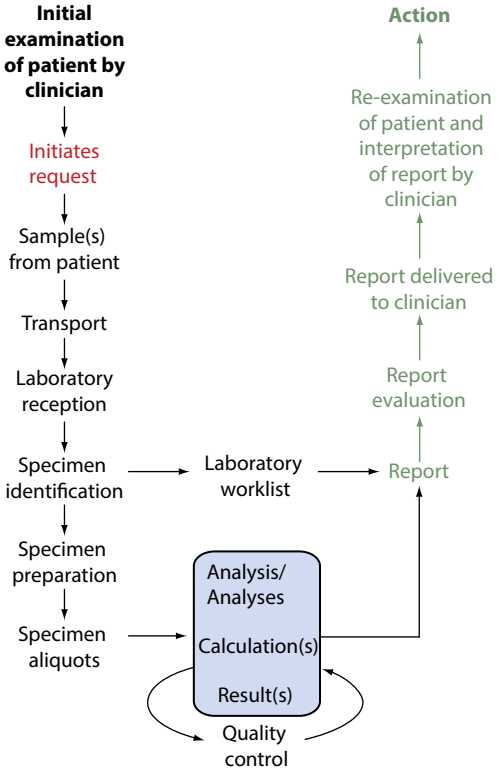


Figure 1.24 Sequence of events involved in obtaining a clinical test result.

Prior to any test being requested by a clinician, careful thought should be given as to whether the test is necessary and how its results will affect the management of the disease to the benefit of the patient. If this is not the case, then one has to consider the value in requesting and performing the test. Unfortunately, clinicians sometimes request clinical tests that are unnecessary and will not be of benefit in treating the patient. This problem often occurs when using forms on which tests can be requested simply by ticking a box. Requesting unnecessary tests poses a number of problems for patients, clinicians, biomedical scientists (medical technologists in the USA) and hospitals. The test means that the patient is inconvenienced, as extra specimens are required. Unnecessary tests can be misleading and result in poor patient management while imposing a financial burden on the hospital. The increased workload for laboratory scientific staff may make the clinician in question rather unpopular! However, set against this is the fact that many of the machines used in the hospital for clinical analysis routinely test for a number of analytes, whether they are asked to or not, since it is easier to set up the machine in this way than to adjust it for individual patients.

Collection, transport and storing of clinical specimens

Clinical specimens are collected into a variety of containers, from tubes to contain simple swabs of the ear, nose and throat to a variety of larger samples (*Figure 1.25*),

test. The screening involves a person collecting a faecal sample at home and then sending it to a laboratory for processing. The test detects the presence of faecal occult blood in the sample and can therefore lead to the identification of patients with early-stage bowel cancer. However, no screening is 100% reliable and some abnormalities may go undetected, hence the importance of regular screening tests.

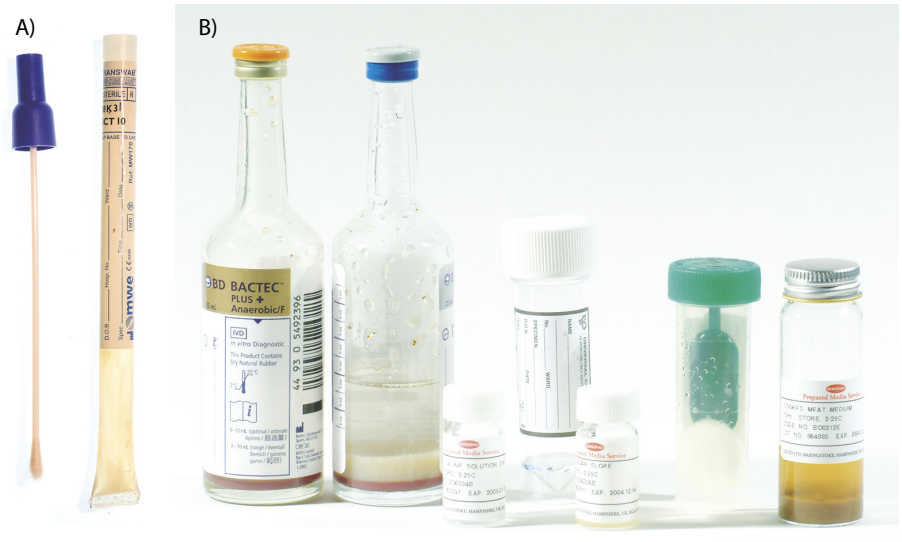


Figure 1.25 (A) A general swab and its transport tube used in taking samples from, for example, ears, nose and throat. (B) Some containers used in the collection of clinical specimens. Courtesy of BUPA Hospital, Manchester, UK.

such as solid **biopsy** material, that is, a small piece of tissue taken from the patient by surgery. Blood can be collected using a simple thumb prick, especially for point-of-care tests (POCTs; *Margin Note 1.4*), or, more usually, using a syringe that delivers it into specialised tubes, a system called vacutainers (*Figure 1.26*). Blood needs to be collected with care (*Chapter 13*). If too many erythrocytes burst, known as **haemolysis**,

Margin Note 1.4
Point-of care testing

Point-of-care testing (POCT) is clinical testing that is performed physically closer to the patient, that is, outside a traditional clinical/pathology hospital laboratory. Conventional clinical laboratory testing involves collecting a specimen from the patient, which is then transported to a central laboratory (often located some distance away) for analysis involving several steps before the results are reported back to the clinician who is caring for the patient. This takes time and can mean patient treatment is delayed. However, POCT uses portable devices and provides rapid test results nearer to where the patient is being cared for or treated. Consequently, it can enable treatment to start earlier, which can result in better clinical and economic outcomes than a traditional laboratory evaluation.

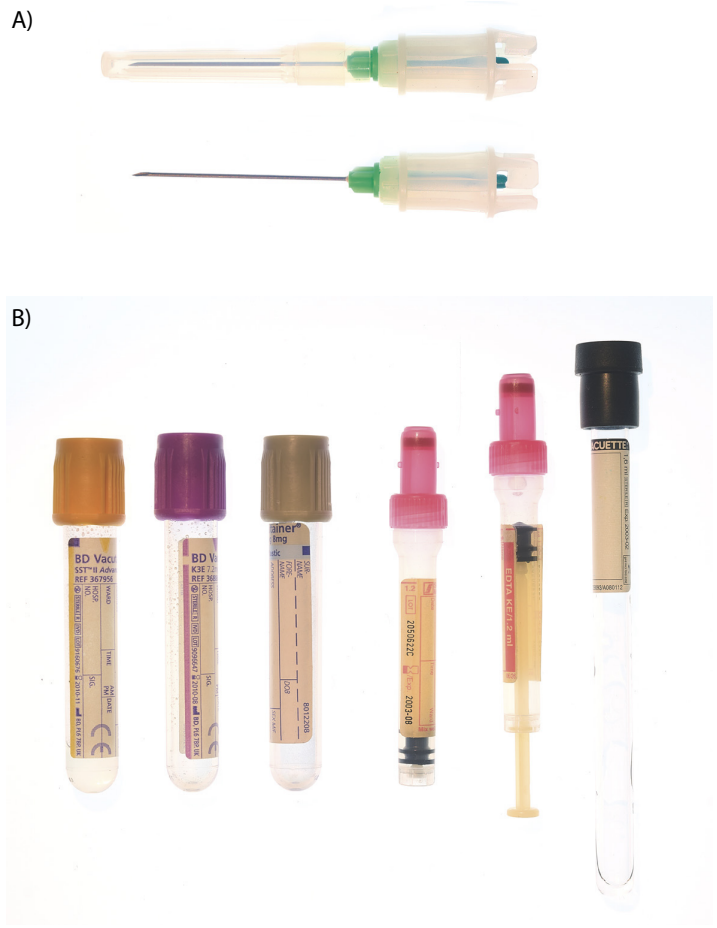


Figure 1.26 (A) Syringe needles and (B) blood collection tubes with colour-coded lids, each type being used for a different clinical test.

the specimen will be unsuitable for the determination of some analytes: for example, the value obtained for 'serum K^+' will not be a true value, since potassium ions are released from haemolysed blood cells. Blood should not be collected from an arm that is receiving an infusion, as this will have diluted the blood. Often, in such cases, the measured concentrations of electrolytes and glucose in the blood samples resemble those of the infusion fluid. For some analytes, the blood must be collected into a tube containing an anticoagulant or preservative. Specimens of blood for glucose determination must be collected into a tube containing fluoride ions (F^-), since this inhibits glycolysis and prevents the utilisation of serum glucose by erythrocytes and leukocytes. Occasionally, blood is collected into the wrong tube and then decanted into the correct tube. This can cause a number of problems. For example, blood collected into a tube containing ethylenediaminetetraacetate (EDTA) will be unsuitable for the determination of serum Ca^{2+}, since EDTA, an anticoagulant, chelates and removes available Ca^{2+} (see *Figure 12.28*).

The transport of specimens and their storage must be considered carefully, since an inappropriate environment can influence the values of clinical test results. Swabs, for example, obtained during a microbiological investigation of an affected site contain only a small volume of specimen and dry easily. They therefore need to be transported to the laboratory as quickly as possible to preserve the pathogens. Blood specimens that have been stored overnight may show erroneously high concentrations of serum K^+, phosphate and activities of erythrocyte enzymes because these all leak from the cells during storage. To prevent this happening, the serum should be separated from blood cells immediately following collection and stored separately if it is to be analysed the following day.

Certain samples need to be collected over a specific period: for example, urine must be collected over a 24-hour period when determining creatinine clearance values (*Chapter 8*) and stools for faecal fat determination to assess malabsorption require collection over a three-day period (*Chapter 11*). The results obtained for such tests often lack accuracy because of the practical difficulties in obtaining accurately timed specimens from the patient.

1.8 EVALUATING LABORATORY TESTS

The clinical tests used in hospital laboratories undergo thorough evaluations prior to use in a laboratory. A number of factors are assessed, including accuracy, precision, reliability, practicality, safety, ease of use, duration and cost.

Analytical variation

Accuracy refers to the ability of a method to give results that are close to the true value of the analyte being measured. The results obtained from laboratory tests may be based on subjective assessment, for example following the microscopic examination of a biopsy tissue during the investigation of a possible malignancy. These types of assessments rely heavily on the experience of the practitioner in recognising and identifying key changes. However, interpreting quantitative data from tests, such as blood amylase activities in acute pancreatitis or the concentration of creatinine in the serum of a patient with suspected renal failure, requires referral to their reference ranges (see *Box 1.1*).

Precision refers to the ability of a method to provide the same result every time it is used. It is assessed by repeatedly measuring samples taken from a single specimen and from batches of samples. The variation in the results may be assessed by calculating statistical parameters, such as their standard deviation or coefficient of variance (see *Box 1.1*).

Analytical sensitivity refers to the ability of a method to detect small amounts of the analyte under investigation (*Box 1.2*). Often, substances in biological materials are present in extremely low concentrations and it may be necessary for the clinical test to detect such low concentrations of analyte and, indeed, monitor changes in its concentration. A term related to analytical sensitivity is the limit of detection, which is the smallest amount of a substance that can be distinguished from the zero value.

Analytical specificity refers to the ability of a method to detect only the test substance (*Box 1.2*). Biological materials contain many components and some of these may interfere with the test being used, giving rise to unreliable results.

BOX 1.2 CLINICAL SPECIFICITY AND SENSITIVITY

The ability of a test to discriminate between diseased and healthy states is described by its clinical **specificity** and **sensitivity**. If a test for a particular disease gives a positive result in affected patients, the result is referred to as a true positive (TP). However, if a positive result is obtained in an individual who does not have the disease, this is referred to as a false positive (FP). In individuals without the disease, the test results should be a true negative (TN) but, occasionally, a negative result is obtained in a patient who has the disease, and this is referred to as a false negative (FN).

The specificity of a test is the measure of the incidence of negative results in individuals free of the disease, defined as:

$$\text{Specificity} = \text{TN} \times 100 / \text{TN} + \text{FP}$$

A test with a specificity of 90% means that, on average, 90 out of 100 individuals without the disease would give a negative test. Conversely, 10 of these individuals would give a positive result even though they do not have any disease.

The sensitivity of a clinical test is a measure of the incidence of positive results in individuals affected by the disease. Sensitivity is expressed as:

$$\text{Sensitivity} = \text{TP} \times 100 / \text{TP} + \text{FN}$$

A test of 90% sensitivity means that, on average, 90% of individuals with the disease will give a positive test while the remaining 10% of individuals with the disease would give a negative result.

Ideally, both the specificity and the sensitivity of a test should be 100%, that is, it should give a negative result in all individuals without disease and a positive result in all patients affected by the disease and would thus effectively discriminate completely between the diseased and healthy states. Unfortunately, such perfection rarely occurs, and tests almost always have some degree of overlap (*Figure 1.27*). Indeed, factors that increase the specificity of a test often decrease its sensitivity, and vice versa.

The ability of a clinical test to detect a disease is influenced by the prevalence of that disease in the population being studied, that is, its predictive values. The predictive value of a positive test is defined as:

$$\text{Predictive value of positive test} = \text{TP} \times 100 / \text{TP} + \text{FP}$$

and a negative test is defined as:

$$\text{Predictive value of negative test} = \text{TN} \times 100 / \text{TN} + \text{FN}$$

The range of values obtained for any test in healthy individuals usually overlaps with those obtained from patients with the disease: some patients who are genuinely ill will give test results that imply they are healthy (FN), whereas others who are not ill appear to have the disease (FP). However, if extreme values are used in the test for comparison, then the number of FN results will be reduced or eliminated. The method will, however, detect more FP results (see *Figure 1.27*). Thus, the test will have high specificity but low sensitivity. If the cut-off value is reduced, then the number of FP results would be reduced but at the expense of increasing the number of FN results. In this case, the test has a high sensitivity, but only by decreasing its specificity (see *Figure 1.27*).

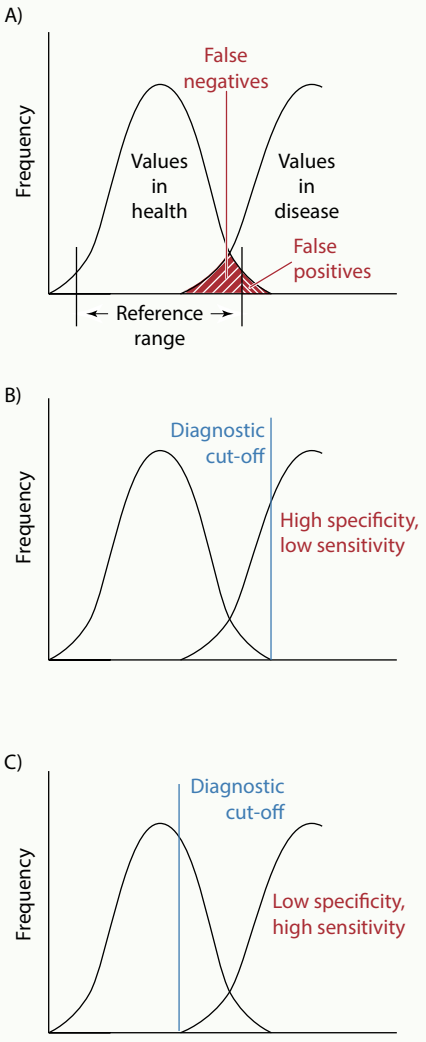


Figure 1.27 (A) The range of results for tests in healthy and diseased populations overlap, and so some patients with disease will have results within the reference range (false negatives), whereas others without disease will fall outside the reference range (false positives). (B) If the diagnostic cut-off value is set too high, then this will reduce false positives but increase the number of false negatives, that is, the test will have high specificity but low sensitivity. (C) If the diagnostic cut-off value is set too low, then the number of false positives increases whereas the number of false negatives decreases, that is, the test has low specificity and high sensitivity.

Whether the sensitivity or the specificity of a test should be increased depends on the disease under investigation and the consequences of making an incorrect diagnosis. Hence, when screening for a disease with severe or fatal consequences, the test must have a high sensitivity. This ensures it will detect all results that are TP, although some FP results will also be detected. Patients with positive results can then be investigated further to identify those with FP results. This test should also have a high predictive value for a negative test so that affected individuals are not missed when screening for the condition. Conversely, in some circumstances, it might be more important to have a test with high specificity. For example, if the purpose of the test is to identify and select patients

for treatment with a new drug, then it is necessary that the test has a high specificity. This will ensure that individuals without the disease are not selected and treated. This type of test should have a high predictive value for a positive result so that the number of individuals with FP results is minimised and they are not subjected to any unnecessary treatment.

Clinicians must be careful in interpreting test results. It is obviously unsatisfactory to tell a patient that he or she is suffering from cancer when this is not the case, and vice versa. The wrong diagnosis may well lead to the wrong treatment being given or, in the case of FPs, no treatment being given for the clinical condition.

Biological variation

In addition to analytical variations, test results are subject to biological variation. The discrimination between normal and abnormal results can be influenced by a variety of biological factors, including:

- the sex of the patient
- the age of the patient
- the patient's diet
- the time of collecting the sample
- the posture adopted by the patient during sample collection
- whether the patient is stressed
- whether the patient has been physically active
- the menstrual state of a female patient
- pregnancy
- whether the patient is taking drugs (legal or illegal)

All influence the results of the test; for example, plasma iron and urate values are higher in male than female patients, and the activity of serum alkaline phosphatase is greater in growing children than in adults. Variations in diet may affect the concentration of certain analytes, such as cholesterol (*Figure 1.28*). The values of some analytes, such as the concentration of cortisol (see *Figure 7.48*) in the plasma, shows a diurnal variation. Blood should be collected from a seated patient, since differences in posture at the time of blood collection can influence the concentrations of a number of analytes, including the concentration of plasma proteins. Stress influences the release of a number of hormones, such as adrenaline and cortisol, while the concentrations of serum analytes, like creatine kinase and lactate, increase following exercise. The concentrations of ovarian hormones are strongly influenced by the menstrual cycle: corticosterone varies by as much as 50% during different stages of the cycle. The nature and concentrations of many hormones also change during pregnancy. Lastly, some drugs can influence results. For example, patients on oestrogen-containing oral contraceptives often have an increased concentration of total plasma protein (*Chapter 7*).

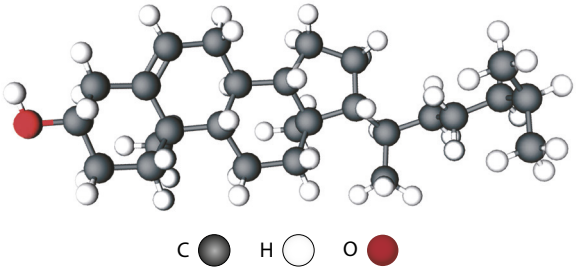


Figure 1.28 Computer drawn model of cholesterol.

Margin Note 1.5
False negatives and false positives

False negative or positive results from a test can have serious consequences (see *Box 1.2*). This applies even more so to tests that patients can carry out at home. One can imagine the joy or anguish, respectively, that a false positive in a pregnancy test might cause. The recent development of a quick home test for HIV–AIDS based on a saliva test is perhaps a more severe example. This test has raised fears that people who obtain a positive result may kill themselves.

1.9 QUALITY OF CLINICAL/PATHOLOGY TEST RESULTS

The results of tests performed in pathology laboratories assist with diagnosing and the monitoring of treatment. Therefore, it is essential to assure the quality of laboratory results. Erroneous results have the potential to cause considerable harm (both physical and psychological) to patients and must be avoided. All laboratories have practices and procedures to ensure erroneous results are minimised and that good quality results are provided.

Errors can arise at the three different stages of analysis:

1. preanalytical
2. analytical
3. postanalytical

Preanalytical errors occur before the sample has been analysed. In general, they result from inappropriate methods of collection or incorrect labelling, handling, transport or storage of specimens.

Analytical errors arise during the laboratory testing procedure. The experimental mistakes that can give rise to analytical errors are detected by introducing systems for each clinical test to warn when errors occur. This is normally achieved by analysing a **control sample** within each batch of tests. A control sample is one that is identical in composition to the test samples except that it contains a known concentration of the test analyte. All samples, including the control sample, must be treated identically. For example, if the concentration of an analyte in serum is being determined, then the control should be serum, not water, containing a known concentration of the analyte in question. The control sample is usually an aliquot from a larger sample for which the mean and SD have already been determined. The results for control samples are usually recorded graphically, so that changes in the quality of the method are detected as soon as they arise. A common chart used for quality control purposes is the Levey–Jennings chart in which the control limits are set at $\bar{x} \pm 2SD$ and $\bar{x} \pm 3SD$. If the control values fall outside the ±2SD limit, that is, there is drift away from the accepted limits, there is only a 5% probability that the result lies in a normal distribution around the mean and is still valid. Any results that lie outside the ±3SD warning limit suggest a problem is occurring with the method (*Figure 1.29*; see also *Box 1.1*). Problems could include unstable reagents in the analyser, defects with its temperature control or contamination, all of which require investigation.

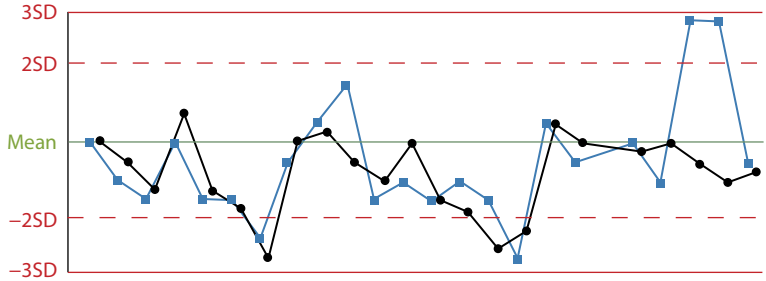


Figure 1.29 A Levey–Jennings chart used to assess quality of quantitative results of tests. The figure shows the values of controls for two tests plotted as means ±2 and 3 standard deviations (SD). When results are plotted this way, any trend affecting the quality of test can be detected. Note the values lying outside the ±2SD limit. However, none lie outside the ±3SD warning limit. Courtesy of Department of Clinical Biochemistry, Manchester Royal Infirmary, UK. ⏎

Postanalytical errors usually arise because of transcriptional errors in reporting the value of a test result. Occasionally, these are gross (and usually obvious), such that they bear no resemblance to values seen in health or disease. Such errors are referred to as 'blunders'. To reduce the number of blunders, results should be checked thoroughly by senior staff before being sent to clinicians.

Quality control and clinical audit

Most laboratories have their own quality control samples and these are used for internal quality assurance purposes. Many countries now participate in external quality assurance, whereby quality control samples are sent to participating laboratories from a central source to assess the analytical performance of their methods for particular analytes. Furthermore, to ensure the quality of service provision, many laboratories follow a set of procedures required for accreditation by external agencies. These procedures ensure good laboratory practice (GLP) and cover all aspects of the laboratory that are involved in the production of test results. These procedures ensure that all laboratory staff are adequately trained and have clearly defined responsibilities. The equipment used should be of adequate standard, with a logbook showing a full record of maintenance and faults. All methods used in the laboratory are standardised, fully documented and appropriate for the analyses. Full details of each method are provided as a standard operating procedure (SOP) that includes details of specimen handling, the analytical method, the equipment used and quality control procedures.

Clinical audit is a process where the practices and procedures involved in patient care are monitored and, if necessary, revised to provide a more efficient and cost-effective service that should ultimately benefit the patient. Many laboratories participate in some form of audit to improve the quality of the services they provide. Audit is part of the process of ensuring quality and is usually divided into five stages (*Figure 1.30*). The first

stage is the observation and review of current practices and procedures in the laboratory. The second stage involves the identification of areas of concern that could be improved. Questions are asked as to whether the current service can be provided more economically. Third, a series of changes are devised to rectify and improve the identified area(s) of concern. Fourth, these changes are implemented and steps are taken to ensure compliance. Finally, in the fifth stage, the changes are monitored and compared with previous procedures to assess whether there is indeed any improvement in the service provided or whether the revised procedures are, indeed, more cost-effective.

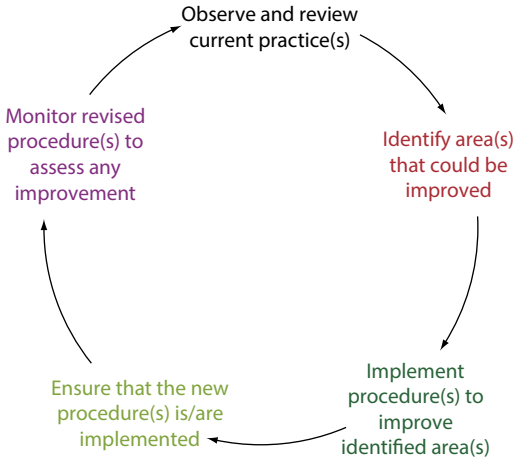

Observe and review current practice(s)
Identify area(s) that could be improved
Implement procedure(s) to improve identified area(s)
Ensure that the new procedure(s) is/are implemented
Monitor revised procedure(s) to assess any improvement

Figure 1.30 Outline of the five stages of a clinical audit.

A clinical audit is usually followed by a re-audit after an appropriate period. The audit may include several processes, such as the initial stages of test requesting, specimen collection and transport. The audit may wish to investigate whether appropriate advice is available to clinicians requesting tests, whether the test request forms are easy to use or whether appropriate containers are provided for specimen collection. Other types of audit processes may relate to the analytical service provided by the laboratory, such as whether the repertoire of tests offered is appropriate to the needs of the clinical service. A clinical audit may wish to investigate whether provision of the laboratory service out of hours is efficient and cost-effective and whether test results are being returned to the clinicians at the right place and within an appropriate time.

1.10 CASE STUDIES

CASE STUDY 1.1

The serum K^+ concentration of Emma, a 22-year-old patient about to undergo surgery, was determined. The laboratory provided a value of 35 mmol/L to the ward. The reference range for serum K^+ is 3.5–5.0 mmol/L.

Questions

(a) What is the most likely explanation of this result?
(b) Discuss its implications.

CASE STUDY 1.2

A new test for detecting prostate cancer has been developed and is undergoing clinical evaluation to determine how effective it is in diagnosing this condition. The study included 100 healthy men and 100 with known prostate cancer. All were screened using the new test, which involved determining the concentration of a serum tumour marker. The patients were rated either positive or negative depending on whether the measured concentration of the tumour marker was above or below a certain cut-off value. The results obtained were as follows:

	Positive	Negative
Patients with prostate cancer	95	5
Healthy individuals	5	95

Questions

(a) What are the sensitivity and the specificity of the test?
(b) Discuss these values.

CASE STUDY 1.3

Ten hospital clinical/pathology laboratories participated in a quality assurance scheme by testing aliquots of a blood specimen for a particular analyte. Their test results are given in *Table 1.3*.

Laboratory	Test value (μmol/L)
1	103
2	100
3	101
4	105
5	102
6	130
7	104
8	104
9	102
10	101
Mean	105

Table 1.3 Results obtained by the 10 clinical/pathology laboratories participating in the quality assurance scheme.

Questions

(a) Comment on these test results.
(b) Should any advice be given to any of the laboratories?

1.11 SUMMARY

Disease refers to any abnormality or failure of the body to function properly. For medical treatment to commence, the disease needs to be diagnosed and its aetiology established. This also helps to define the prognosis, that is, the likely course of the disease and its outcome. Diseases have a number of possible causes. They may be exogenous, such as infections or trauma (accidents), or they may be endogenous, such as diabetes or cancer. Some diseases may be caused by a range of factors and are called multifactorial and, for some, the cause may be unknown, and these are referred to as idiopathic. Pathogenesis describes how the aetiological agent produces the clinical signs and symptoms and the pathological changes characteristic of that disease. This enables the clinician to make a diagnosis and prescribe treatment. A prognosis should also emerge at this stage.

Diseases may be classified into a number of types: degenerative, endocrine, genetic, homeostatic, iatrogenic immunological, infectious, neoplastic, nutritional, psychogenic, toxic and traumatic. Epidemiology is the study of how diseases spread in populations. This is of importance in the control of diseases.

Various types of laboratories specialise in investigating the pathology of diseases. A laboratory may measure the concentrations of analytes in blood, serum and urine; identify infectious agents such as viruses and bacteria; characterise genetic diseases by examining the patient's chromosomes; or identify problems with the blood, for example those involving defective haemoglobin or clotting factors. The laboratories report their findings to the clinician treating the patient to help in diagnosis, determine treatment(s) and assess how well any treatment is working. Pathology laboratories must work to high standards of accuracy, otherwise the clinician and patient may be misinformed about the disease and its prognosis and the wrong medical care and clinical management given. Therefore, pathology laboratories take steps to standardise and check their procedures daily and investigate new tests exhaustively before they are introduced.

QUESTIONS

1. Which of the following terms best describes a test used to detect disease before it presents clinically?
   (a) Diagnostic test.
   (b) Sensitive test.
   (c) Screening test.
   (d) Prognostic test.
   (e) Specific test.

2. Which of the following terms is used to describe the cause of a disease?
   (a) Manifestation.
   (b) Aetiology.
   (c) Pathogenesis.
   (d) Mortality.
   (e) Epidemiology.

3. A Na^+ solution with a concentration of 10 mmol/L was investigated by two methods: X and Y. The values obtained with methods X and Y were 11.8 mmol/L and 9.7 mmol/L, respectively. Which ONE of the following statements is correct?
   (a) Method X is more accurate than method Y.
   (b) Method X is more precise than method Y.
   (c) Method Y is more accurate than method X.
   (d) Method Y is more precise than method X.
   (e) Method X is as accurate as method Y.

4. The concentration of Mg^{2+} in the serum of an individual is 0.5 mmol/L. The reference range is 0.7–1.0 mmol/L.
   (a) Is it possible to be 100% certain that this individual has a deficiency of Mg^{2+}?
   (b) Explain your answer.

5. Indicate which of the following statement(s) is/are true.
   (a) Pathology is the scientific study of diseases.
   (b) Health may be defined as an absence of signs and symptoms associated with a disease.
   (c) The aetiology, pathogenesis, clinical manifestations and prognosis of a disease or clinical condition allows it to be categorised and diagnosed.
   (d) Prions are native proteins that cause spongiform encephalopathies and are endogenous and exogenous causes of disease.
   (e) The incubation period is the time that must elapse before a disease becomes apparent following exposure to its aetiological agent.
   (f) Syndromes are characterised by multiple clinical abnormalities that together form a distinct clinical manifestation.
   (g) Congenital diseases are inherited conditions that are present at birth.
   (h) Epidemiological studies can provide data about, for example, the prevalence, incidence, signs and symptoms of a disease.
   (i) Most laboratories use healthy blood donors as a representative sample of the local population when determining the reference range for a particular analyte.
   (j) Levey–Jennings charts are essential tools in performing a clinical audit.

FURTHER READING

Ahmed N *et al.* (2022) *Biomedical Science Practice*, 3rd edn. Oxford University Press.

Badrick T (2021) Integrating quality control and external quality assurance. *Clin Biochem* **95**:15–27.

Bolboacă SD (2019) Medical diagnostic tests: a review of test anatomy, phases, and statistical treatment of data. *Comput Math Methods Med* **2019**:1–22.

Doyle K and Bunch DR (2023) Reference intervals: past, present, and future. *Crit Rev Clin Lab Sci* **60**:466–482.

Erasmus RT and Zemlin AE (2009) Clinical audit in the laboratory. *J Clin Pathol* **62**:593–597.

Fricker Jr RD and Rigdon SE (2020) Investigating a disease outbreak. *Signif (Oxf)* **17**:14. doi: 10.1111/1740-9713.01372.

Henny J and Hyltoft Petersen PH (2004) Reference values: from philosophy to a tool for laboratory medicine. *Clin Chem Lab Med* **42**:686–691.

Irjala KM and Grönroos PE (1998) Preanalytical and analytical factors affecting laboratory results. *Ann Med* **30**:267–272.

Lau B *et al.* (2020) Perspectives on the future of epidemiology: a framework for training. *Am J Epidemiol* **189**:634–639.

Sinclair G *et al.* (2023) EQA oversight and escalation. *Biomed Sci* (Dec 2023) 34–36.

Stewart A (2022) *Basic Statistics and Epidemiology: A Practical Guide*, 5th edn. CRC Press.

Weinstein S *et al.* (2005) Clinical evaluation of diagnostic tests. *AJR Am J Roentgenol* **184**:14–19.

Taylor & Francis
Taylor & Francis Group
http://taylorandfrancis.com

Pathogens and parasites

2

OBJECTIVES

After studying this chapter you should be able to:

- define pathogen, parasite and related terms
- indicate the general course of an infection
- list the types of pathogens and parasites
- summarise the actions of microbial virulence factors
- outline examples of diseases caused by pathogens and parasites

2.1 INTRODUCTION

Pathogens are traditionally defined as microorganisms that can cause infectious (from the Latin *inficere*, 'to put in') diseases. Parasites are organisms that live on or in patients (hosts) using his or her biological resources and so cause them harm. Neither definition, however, is particularly exclusive and, indeed, there is considerable overlap between the two groups of agents. For example, prions are aberrant proteins but can cause infections. Viruses are not microorganisms but some are pathogenic and show many features that resemble those of a parasite in the way they obtain materials and energy from host cells. Bacteria, single-celled fungi and protozoa more closely conform to the definition of a pathogen given above. In contrast, 'classic' human parasites, such as some macroscopic fungi and helminth worms, display many of the traits of a pathogen. Hence, the terms pathogen and parasite are virtually interchangeable.

Communicable diseases, for example acquired immunodeficiency syndrome (AIDS), influenza, malaria and poliomyelitis, are transmitted from one patient to another individual. In general, the pathogen/parasite concerned cannot remain viable for extended periods outside its host.

Diseases that can be transmitted directly from a patient to another human are referred to as **contagious diseases**, and the patient is called the definitive host. **Pathogenesis** may be defined as the molecular and biochemical mechanisms that allow pathogens to cause diseases. Many pathogens/parasites infect intermediate hosts as part of their transmission cycle (*Section 2.3*). In addition to such vectors, some infectious diseases are spread by mechanical ones on, for example, the bodies of cockroaches and houseflies (*Figure 2.1*) or in contaminated food and water.

However, many previously prevalent diseases have been controlled by advances in social hygiene, such as potable water supplies and the efficient disposal of sewage and other wastes (*Figure 2.2*), increased standards of living, and the development of successful vaccines (*Chapter 3*). Also, drugs such as antibiotics and monoclonal antibodies (see *Box 4.5* and *Chapters 3* and *17*) are available to treat infectious diseases and, in many cases, will effect a cure. It is, however, possible for drug regimens to exacerbate the problem, since they can remove commensals, allowing antibiotic-resistant pathogens to flourish. Indeed, numerous pathogens are now resistant to many drugs (*Chapter 3*). An adequate immune response is often necessary to combat disease, since drugs alone may fail to eliminate the infection. Thus, all pathogens and parasites must overcome the defence systems present in their hosts (*Chapter 4*).

Unfortunately, the emergence of new pathogens and the development of resistance to current treatments by existing pathogens and parasites means that infectious diseases will probably always be with us.

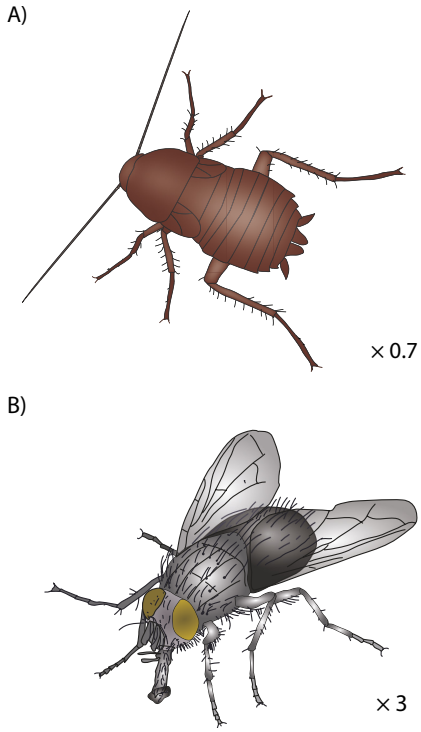


Figure 2.1 (A) Cockroach (*Periplaneta americana*) and (B) housefly (*Musca domestica*).

DOI: 10.1201/9781003296836-2

Figure 2.2 Modern (A) water treatment (purification) facility (Williamsgate Water Treatment Works) and (B) activated sludge (sewage) wastewater treatment plant (Davyhulme Wastewater Treatment Works). Courtesy of United Utilities, UK.

2.2 INFECTIOUS DISEASES

Infectious diseases may be caused by agents that are molecular in structure, such as prions and viruses; single-celled prokaryotic bacteria, eukaryotic protozoa and yeasts; or multicellular organisms, such as some fungi, certain worms (generally referred to as **helminths**, for example nematodes and flukes) and arthropods (such as mites).

The body is exposed to many pathogenic viruses, microorganisms and multicellular parasites, although most microorganisms found associated with the body are harmless and live as commensals or symbionts. Unfortunately, it is likely that any microorganism with the ability to live in or on humans will sometimes become an **opportunistic pathogen**, especially if the balance between the usual microorganisms present and the immune system is disturbed. For example, wounds can become badly infected with bacteria that normally live harmlessly on the skin and bacteria found in the gastrointestinal tract (GIT) can cause peritonitis if the gut contents enter the peritoneum (*Chapter 11*), while some yeasts can severely infect immunocompromised patients. In general, such infections are not transferable to other healthy humans. However, some pathogens regularly cause diseases while others do not. For example, *Pseudomonas aeruginosa* (*Figure 2.3*) can cause overwhelming disease in patients whose immune system is compromised but not in those with intact defences.

The success of pathogens and parasites depends on their ability to colonise host tissues and to counter the host defence mechanisms. Virulence is measured by the **infective dose** and the severity of the disease caused. For example, as few as 10–100 *Shigella dysenteriae* cells can cause shigellosis, but more than 10 000 cells are needed of the less virulent salmonella or cholera bacteria. True pathogens are equipped with a range of **virulence factors** (*Section 2.6*). The strains of some bacterial species, such as pathogenic

BOX 2.1 INFECTIOUS DISEASES IN ANTIQUITY

The earliest human skeletons show evidence of a variety of infectious diseases. Evidence from mummies preserved in ancient Egypt 4000 years ago shows that the population suffered from diseases such as tuberculosis, trachoma and dental caries. The Greek historian Thucydides (c. 460 to c. 400 BCE) described an outbreak of 'plague' in the city of Athens during its siege by Sparta and its allies during the Peloponnesian war of 430–427 BCE. It is estimated to have killed 25–35% of a population thought to number 250 000–300 000. Not perhaps surprisingly, Sparta won the war. The identity of the disease responsible is uncertain. Cholera, Ebola, influenza, measles, smallpox and typhoid fever have all been suggested, but it is unlikely that the culprit will ever be unambiguously identified. Hence, infectious diseases have always been a threat, but their intensity increased with the development of urban life, with its increased population densities.

Margin Note 2.1 A summary of some relevant terms

Communicable diseases, for example acquired immunodeficiency syndrome (AIDS), influenza, malaria and poliomyelitis, are transmitted from one patient to another individual. In general, the pathogen/parasite concerned cannot remain viable for extended periods outside its host.
Noncommunicable diseases cannot be transmitted from one patient to another individual. Examples include cancers, cardiovascular conditions, diabetes and inborn errors of metabolism.
Infectious diseases is another term for communicable diseases. Note, however, that **contagious diseases** are a discrete category of infections that can be transmitted only by direct contact between individuals.
Carriers are humans who harbour a pathogen but are asymptomatic, although they can infect others.
Zoonoses are infections that can be transmitted from an animal directly to a human.

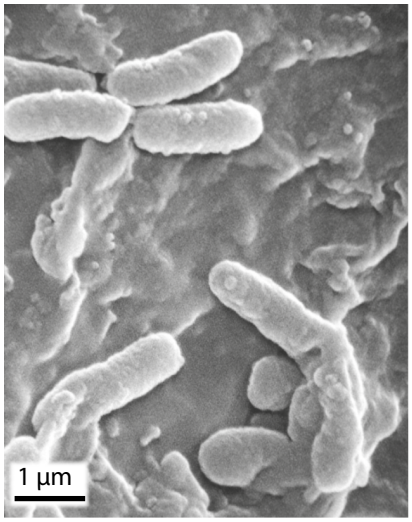


Figure 2.3 Electron micrograph of *Pseudomonas aeruginosa*. Courtesy of the Centers for Disease Control and Prevention/ Dr J. Carr.

forms of *Escherichia coli* (*Figure 2.4*), can produce different virulence factors that cause, for example, diarrhoea, urinary tract infections or sepsis. Other strains, however, do not produce virulence factors or do so to a lesser extent and are therefore not pathogenic, except when they infect an immunocompromised host.

Protozoa, fungi and helminths are eukaryotic organisms. They therefore share many biochemical features with humans. Also, they have many seemingly sophisticated ways of countering the host immune system, which often makes them difficult to treat or eradicate when they cause disease. Fungi, protozoa and helminth parasites are responsible for a large proportion of infective diseases, particularly in the developing world.

A pathogen must be transmitted from a source to the patient. Direct contact between hosts is the most obvious form of transmission, but coughs and sneezes (aerosols), food, water, and arthropod and vertebrate vectors are all used by various pathogens. The long-term survival of pathogenic microorganisms also depends on them maintaining their infectivity during transmission from host to host.

Figure 2.4 Electron micrograph of dividing *Escherichia coli* cells. Courtesy of Dr A. Curry, Manchester Royal Infirmary, UK.

2.3 COURSE OF AN INFECTION

Virulent pathogens are those well adapted to cause an infection. The course of such an infection, often referred to as its **transmission cycle**, can be considered to follow up to four major stages, namely:

1. adhesion
2. entry
3. localised infection
4. generalised infection

Adhesion is the first stage of infection, given that it prevents the pathogen being swept away and eliminated from the body. Most pathogens are adapted to adhere to cells, usually epithelial, that line the site of entry. Pathogens then enter the body through one of several so-called **portals of entry**. These include the skin and conjunctiva, respiratory tract, GIT, urogenital system and, in the case of fetuses, placenta. Once entry has been gained, conditions for growth, such as temperature and nutrients, must be favourable, but the pathogen must also be able to overcome the local defences. Pathogens are often adapted to infect their hosts through a single portal of entry and do not cause infectious diseases if they enter through a different one.

A localised infection acts as a focus of infection and, indeed, many remain local in nature or are prevented from spreading by host defences. If the pathogenic organisms penetrate tissues and reach the blood or lymphatic systems or enter cells such as phagocytes, they can be distributed throughout the body and infect other tissues and organs causing a generalised infection. Again, certain pathogens are particularly well adapted to infect specific target tissues and organs. During infections, the pathogen multiplies and can be shed from the host into the environment through **portals of exit**. In localised infections, these are the same as those of entry; in a generalised infection, other portals are used. These may include loss from the skin in pus, scales and blood; from the respiratory tract as droplets and aerosols of mucus and saliva; from the GIT in faeces and bile; from the urogenital system in urine, mucus and genital secretions; and, in pregnancy, from the placenta with direct transfer to the fetus.

A pathogen may be eliminated at any stage of the infective process. In some cases, the growth may be held in check but not eliminated. Such latent infections can be activated later, giving recurring infections. In some cases, individuals can recover from a disease, but the pathogen may remain in the body for considerable periods. Such people are symptomless carriers and are reservoirs of the disease.

2.4 DEATHS FROM INFECTIONS

Infectious diseases are the major cause of deaths globally: they are thought to kill 15 million people annually. Deaths associated with human immunodeficiency viruses (HIVs) have reached nearly 40 000 000 since it was first reported in the early 1980s. Tuberculosis (TB), caused by the bacterium *Mycobacterium tuberculosis*, is the 10th leading cause of death globally. A quarter of the world's population is thought to be infected with TB bacteria. In 2023, *M. tuberculosis* infected 10.8 million new patients and killed 1.25 million people. Children were its major victims, with a million becoming ill and approximately 250 000 dying. Deaths due to protozoal infections, such as malaria and Chagas disease, are numbered in the 100 000s annually. Fungal diseases, largely due to infection of *Aspergillus*, *Candida*, *Cryptococcus* and *Pneumocystis* spp., help kill more than 1 500 000 yearly. Parasitic worms, such as species of cestodes, nematodes and trematodes, infect

two billion people worldwide. Helminthiases are often thought to be confined to developing tropical countries, but some types also occur in developed temperate regions. Many of these conditions are not usually fatal but can cause severe ill health by depriving patients of nutrients (*Chapter 10*), damaging tissues and causing loss of blood, and secreting toxins.

2.5 PRIONS

Prion diseases or spongiform encephalopathies include Creutzfeldt–Jakob diseases (CJDs), Gerstmann–Sträussler–Scheinker syndrome (GSS) and fatal familial insomnia (FFI). These diseases generally develop slowly over 10–20 years in older individuals. These rare diseases commonly have incidences of one in a million. Prion diseases occur sporadically or can be familial, that is, genetic in origin (*Chapter 15*), or acquired, that is, resulting from an infection or be iatrogenic in origin.

Prions are degenerate host proteins. The normal form of the protein adopts a largely α helical conformation that is harmless but can refold to a β sheet-rich form that is a pathological conformation (see *Figure 15.18*). A misfolded prion protein can induce a conformational change in a native α prion protein to produce a β type conformation. This new misfolded protein, in turn, can catalyse the conformational change in other native proteins. Eventually this chain reaction produces aggregates of prions in the brain that lead to holes developing in its tissue and a lethal spongiform condition.

The sporadic forms of the diseases occur in individuals with mutations in the prion gene that predispose them to produce the misfolded form of the protein (see *Box 15.1*). However, if misfolded prions enter the body, for example in the diet, they resist digestion and induce disease. They may also enter by iatrogenic means, through surgery or blood transfusions for instance, and can initiate the infectious form of the disease. Symptoms of spongiform encephalopathies are associated with brain damage: loss of memory and intellectual abilities, slurred speech, spastic movements and reduced mobility, and blindness. In addition to this spectrum of symptoms, diagnosis can be aided by scans of the brain to show the prion-associated damage (see *Figure 15.19*) or by analysing samples of cerebrospinal fluid (CSF). All spongiform encephalopathies are incurable and are invariably fatal.

Animals are also susceptible to prion diseases. In the 1980s and early 1990s, scrapie, a prion disease of sheep, was able to jump species and infect cattle given fodder containing sheep protein contaminated with the scrapie prion. The symptoms included a characteristic loss of balance and unsteady gait, leading to the common name 'mad cow disease', which affected several other species, for example elk and mink. Much more unfortunately, contaminated beef entered the human food chain and the prion again jumped species to humans. This produced a particularly insidious spongiform encephalopathy called variant Creutzfeldt–Jakob disease (vCJD; *Figure 2.5*), because it affected a teenager to 40-year-old population, rather than the 65+-year-old group of the sporadic form. Also, the effects became apparent within only a few years. Fortunately, cases appear to have ceased to occur (*Figure 2.6*), although it is possible that those most susceptible have already succumbed and a reservoir of more resistant individuals are incubating the disease; time will tell.

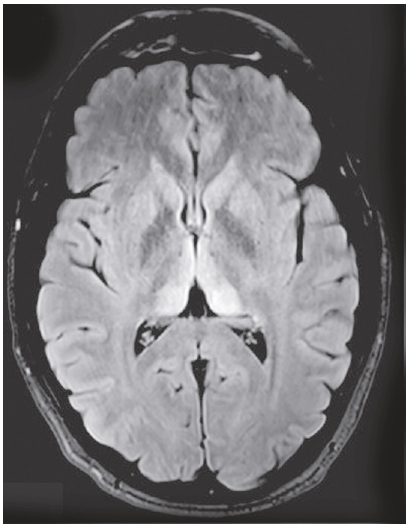


Figure 2.5 Magnetic resonance image of the brain of a patient with variant Creutzfeldt–Jakob disease (vCJD). Courtesy of Dr P. Rudge, The National Prion Clinic, UK. See also *Figure 15.19*.

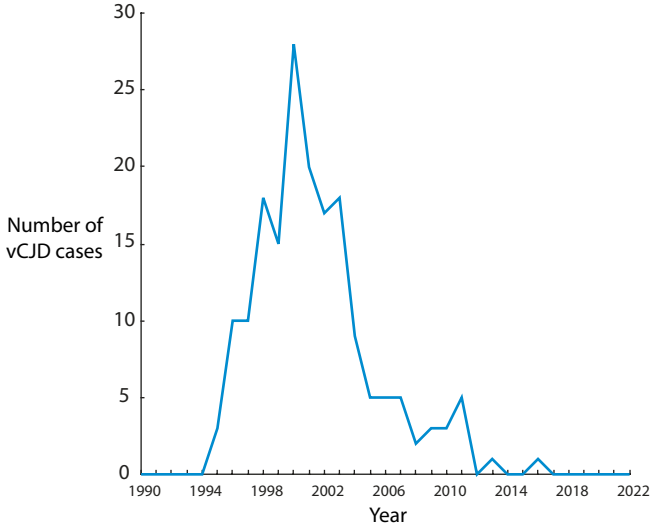


Figure 2.6 The numbers of variant Creutzfeldt–Jakob disease (vCJD) patients per calendar year in the UK. Data from www.cjd.ed.ac.uk.

2.6 VIRUSES

Viruses are obligate intracellular parasites. Virus particles or virions are molecular complexes of nucleic acid and proteins forming a **capsid**. Some viral particles also have a lipid membrane or **envelope** (*Figure 2.7*). Virus particles can be examined by electron microscopy; however, because many can be crystallised, they are increasingly studied by combining X-ray diffraction and molecular modelling methods (*Figure 2.8*).

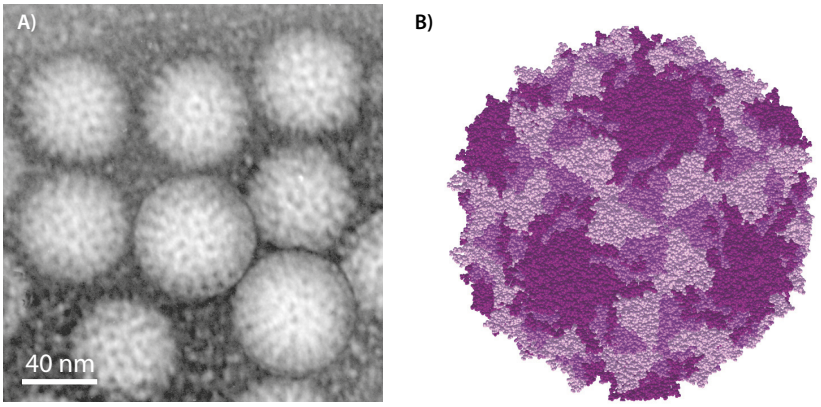


Figure 2.8 (A) Electron micrograph of rotavirus particles. Worldwide, these viruses are the most common cause of severe diarrhoeal illness in children. Courtesy of H. Cotterill, Manchester Royal Infirmary, UK. (B) Molecular model of the capsid of a poliovirus. This group of viruses once caused extensive cases of poliomyelitis, but these have been largely eradicated by widespread vaccination programmes. PDB file 5O5P.

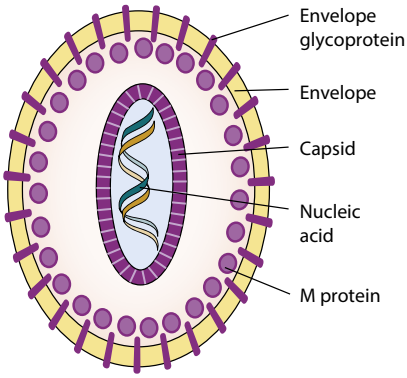


Figure 2.7 Schematic illustrating the structure of a typical virus particle (virion).

The nucleic acid constitutes the genome of the virus. It may be RNA *or* DNA, never both, and, depending upon the type of virus, can be single or double stranded (ss or ds, respectively). The nucleic acid is generally only a few thousand nucleotides or base pairs long. In general, viral genomes contain fewer than about 20 genes, although larger viruses, such as *Vaccinia*, have more than 100. Capsids are formed of proteins called **capsomeres**. Each capsomere is constructed from one or more **protomer** proteins, which may be different or identical polypeptides. The capsid protects the nucleic acid genome and, in nonenveloped viruses, recognises receptors on the surfaces of host cells. In nonenveloped viruses, capsids possess antigenic sites that can stimulate specific immune responses (*Chapter 4*). The envelopes of enveloped viruses are portions of the host cell plasma membrane acquired when the virion leaves its host (see *Figures 1.5, 2.7* and *2.12*). Viral glycoproteins in the envelope, often referred to as 'spikes', are antigenic and also able to elicit immune system responses. Spike proteins function by recognising and binding to specific receptors on target cells. The envelope is generally lined with a layer of matrix proteins. These are multifunctional and, among other roles, inhibit the transcription of host cell genes while helping to regulate the replication and transcription of the viral genome. They also facilitate the assembly of virions and their release from the host cell.

Virus nomenclature and classification

The naming and classification of viruses is overseen formally by the International Code of Virus Classification and Nomenclature (ICVCN) and the International Committee on Taxonomy of Viruses (ICTV), respectively. In general, the nomenclature for any particular virus does not follow any fixed pattern and could be based on:

- the associated disease: polioviruses (poliomyelitis)
- the geographical location where it was initially found: Norwalk virus (Norwalk, Ohio, USA)
- the tissue or organ infected: hepatitis A virus (liver)
- various morphological features of the virion: picornaviruses (*pico*, 'small'; *rna*, RNA)
- acronyms involving features of the virus: ECHO (enteric cytopathic human orphan) virus
- the name of the first person(s) to isolate the virus: Epstein–Barr viruses

The ICTV formally classifies viruses using a hierarchical system of 15 taxa, which resemble those used in classifying organisms, and is based largely on evolutionary relationships. However, species names are still largely irregular, although the aim is to eventually establish Linnean binomials (*Tables 2.1* and *2.2*).

Taxon
Realm
Subrealm
Kingdom
Subkingdom
Phylum
Subphylum
Class
Subclass
Order
Suborder
Family
Subfamily
Genus
Subgenus
Species

Table 2.1 The 15 taxa of the International Committee on Taxonomy of Viruses (ICTV) hierarchical system used formally to classify viruses.

Taxon	Examples	
	Poliovirus 3	**COVID-19[a] virus**
Realm	*Riboviria*	*Riboviria*
Kingdom	*Orthornavirae*	*Orthornavirae*
Phylum	*Pisuviricota*	*Pisuviricota*
Class	*Pisoniviricetes*	*Pisoniviricetes*
Order	*Picornavirales*	*Nidovirales*
Family	*Picornaviridae*	*Coronaviridae*
Genus	Enterovirus	*Coronavirinae*
Species	Poliovirus 3	SARS-CoV-2[b]

[a] Coronavirus disease 2019. [b] Severe acute respiratory syndrome coronavirus 2.

Table 2.2 Abbreviated examples of the International Committee on Taxonomy of Viruses (ICTV) classification.

Prior to the establishment of the ICTV formal classification scheme, David Baltimore (1938–) in 1971 had established a working classification of viruses based on the natures of their nucleic acids. This scheme is used in parallel with the formal ICTV taxonomy and, indeed, was partially incorporated into it. Baltimore initially arranged all viruses into six, later adding a seventh, disconnected groups without any subdivisions. These groups, or Baltimore classes (BCs), are defined by the nucleic acid present: DNA or RNA and whether it is ds or ss. In viruses with ss RNA genomes, the polarity of the RNA is also a criterion. RNA strands that can be directly translated by host ribosomes are termed positive or plus (+); negative strands (–), however, need their complementary + strand biosynthesised before translation can proceed. Replication also helps define a BC. Specifically, a DNA intermediate must be synthesised during the replication of ssRNA viruses, whereas an RNA intermediate needs to be formed when a dsDNA genome is replicated. The seven BCs are summarised in *Table 2.3*.

Margin Note 2.2
Nobel prize for virus studies

David Baltimore (1938–), Renato Dulbecco (1914–2012) and Howard Temin (1934–1994) all shared the 1975 Nobel Prize in Physiology or Medicine 'for their discoveries concerning the interaction between tumour viruses and the genetic material of the cell'.

Baltimore class	Nucleic acid present
I	dsDNA
II	ssDNA
III	dsRNA
IV	+ ssRNA
V	– ssRNA
VI	ssRNA with a DNA intermediate involved in replication
VII	dsDNA with an RNA intermediate involved in replication

Table 2.3 The Baltimore classes (BCs) of viruses.

Viral infections

Viruses must enter a target or host cell to be replicated (*Figure 2.9*). Normally, viruses infect only specific types of cells, a phenomenon referred to as **tissue tropism**. They bind to the target cell by attaching to specific proteins or carbohydrates on the cell surface (*Figure 2.10*). For example, HIVs attach to a CD4 protein found on certain T lymphocytes and macrophages of the immune system (*Chapter 4*). Other examples of viruses and their cellular receptors are given in *Table 2.4*.

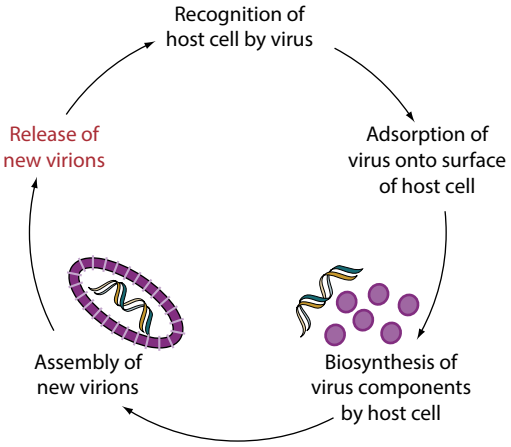

Recognition of host cell by virus
Adsorption of virus onto surface of host cell
Biosynthesis of virus components by host cell
Assembly of new virions
Release of new virions

Figure 2.9 Simplified outline of a viral replication (infection) cycle. See also *Figure 3.21.*

Virus	Cell membrane protein used as virus receptor
Human immunodeficiency viruses (HIV-1 and -2)	CD4 (*Chapter 4*)
Human rhinovirus 91	ICAM I[a]
Sindbis virus	High-density laminin receptor
Coxsackie A	ICAM I[a]
Human coronavirus 229E	Aminopeptidase N
Hepatitis B virus (HBV)	HBV binding factor (a metalloprotease)

[a] Intracellular adhesion molecule.

Table 2.4 Examples of plasma membrane protein receptors to which viruses bind prior to insertion into the cell.

Following infection, the host cell provides the materials, the biological energy and some of the enzyme activities necessary to manufacture new viral particles. In some cases, replication may include an inactive latent state. For example, the virus *Varicella zoster*, which causes chicken pox, enters nerve cells and, after the initial infection, remains dormant. If, however, the host immune system becomes weakened, *Varicella* can become active and cause painful attacks of shingles in the area served by that nerve (*Figure 2.11*). Effective vaccines (*Chapter 3*) against chicken pox and shingles are available; the latter is only suitable for those aged 65 years and over. Some viruses that exhibit latency, with or without sporadic reactivation, are shown in *Table 2.5*.

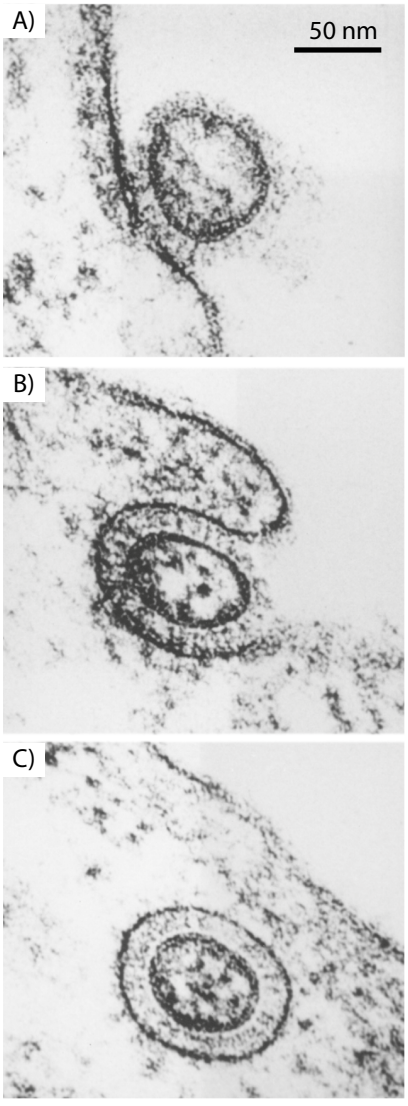


Figure 2.10 Electron micrographs showing the (A) binding, (B) enveloping and (C) internalisation of an influenza virus. Courtesy of Dr J.J. Skehel, National Institute for Medical Research, London. See also *Figure 2.14.*

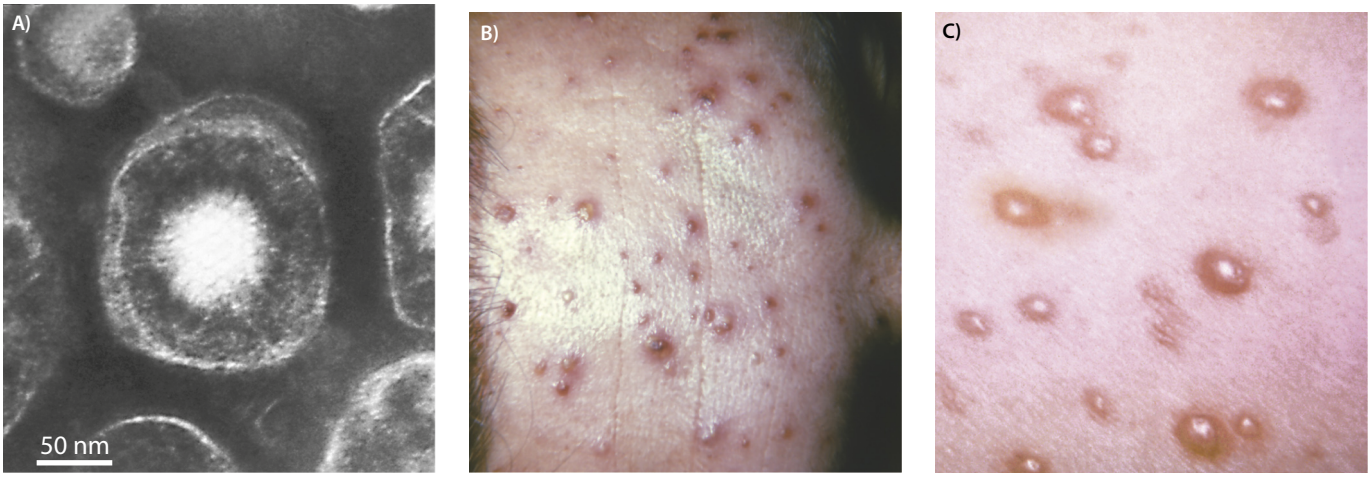


Figure 2.11 (A) Electron micrograph of *Varicella zoster* (HHV-3; see *Table 2.5*). (B) The typical maculopapular rash of chicken pox. (C) A pustular–vesicular rash associated with shingles. Courtesy of the Centers for Disease Control and Prevention/Drs E. Palmer and K.L. Hermann.

Microorganism	Disease
Herpes simplex virus types 1 and 2 (HSV-1 and HSV-2)	Oral and genital herpes
Epstein–Barr virus	Glandular fever
Hepatitis B virus	Hepatitis, hepatocellular carcinoma
Varicella zoster (HHV-3)	Chicken pox, shingles
Measles virus	Measles, subacute sclerosing panencephalitis

HHV, human herpes virus.

Table 2.5 Examples of viruses that exhibit latency. See *Table 3.1* for herpes virus nomenclature. ⏎

The genome of some viral nucleic acids can be integrated into that of the host, eventually causing the infectious disease, as in the case of the AIDS viruses (see *Box 3.3*), or lead to cell transformation and the formation of cancers (*Chapter 17*). For example, the hepatitis B virus can contribute to primary hepatocellular carcinoma development, while certain strains of the human papillomavirus that cause genital warts may contribute to the development of cervical carcinoma.

Viral epidemics and pandemics

Epidemics are increases in the number of patients affected by diseases that are above those normally expected for a population, while pandemics affect many countries and larger populations. Viruses, like bacteria (see *Boxes 2.1* and *2.2*), are responsible for numerous epidemics and pandemics. The AIDS pandemic caused by HIV (see *Figures 1.4* and *2.12*), which has killed nearly 40 000 000 people and caused considerable social upheaval, is discussed in *Chapter 3*. Smallpox (because it has been eliminated as a disease), influenza viruses (which seemingly will be with us forever) and several virus-associated pandemics that have occurred in recent years are discussed below.

Smallpox

Smallpox was for many centuries a scourge. Egyptian mummies as old as 3000 years show smallpox-like lesions. The disease is caused by the variola virus and occurs in two forms. Variola minor produces relatively minor symptoms with a fatality of less than 1%. In contrast, variola major causes fever, headaches, muscle pains and, in children, often convulsions and vomiting 12 days after infection. After a further three to five days, the characteristic papules, vesicles and pustules develop on the face, including the mouth, and the palms and soles of the hands and feet. These detach relatively quickly to leave the distinctive pockmarks. Mortality is approximately 30% following toxaemia and extreme haemorrhaging.

The Antonine Plague (165–180 CE), also known as the Plague of Galen after the Greek physician Galen (129 to c. 216 CE; *Figure 2.13*) who witnessed the outbreak and described the symptoms of the disease and the course of the outbreak, is thought to have been caused by smallpox, although the viral disease measles is also a contender. The Antonine Plague is estimated to have killed 5–10 million people, about 10% of the population of the Roman Empire, including its great philosopher-emperor Marcus Aurelius (121–180 CE), the old emperor in the epic film *Gladiator*. Outbreaks of smallpox also occurred in the Roman Empire between 249 and 270 CE, killing, it is thought, about 5000 people daily in Rome alone, and in Japan during 735–737 CE, which is estimated to have killed approximately 30% of the population. The latter was probably a virgin soil epidemic, that is, one in which a population has not had any previous contact with the disease in question and therefore has little or no immunological defence.

In the fifteenth century, Portuguese explorers and occupiers spread smallpox to parts of western Africa. The discovery of the so-called New World allowed Europeans and captured slaves from West Africa to spread the virus into the Caribbean, then Central and South America in the sixteenth century. Again, the lack of indigenous immunological defence was devastating, with some authors estimating that 90% of the populations succumbed to smallpox and other diseases, such as measles. In the seventeenth century, European settlers brought smallpox into North America, with the deaths of 50% or more of native tribe populations. The virgin soil epidemic scenario was again, unfortunately,

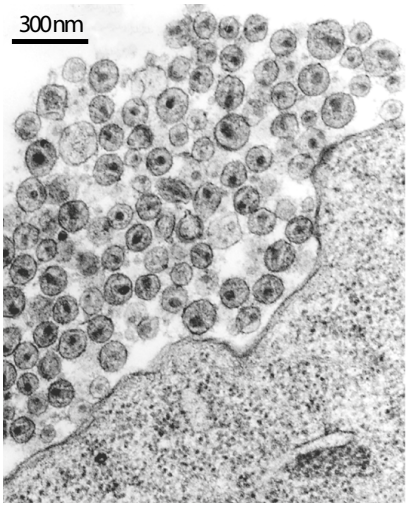


Figure 2.12 Virions of human immunodeficiency virus (HIV) being released from the surface of a human cell. Courtesy of H. Cotterill, Manchester Royal Infirmary, UK. ⏎

Figure 2.13 Galen (in Greek, Galenos; in Latin, Galenus) was the most famous doctor of the Roman Empire. His theories influenced medical practice in the Middle East and Europe for 1500 years. ⏎

repeated among the First Peoples of Australia when British explorers arrived in the eighteenth century.

In 1959, the World Health Organization (WHO) began a global plan to eradicate smallpox with an intensive vaccination programme (*Chapter 3*). Unfortunately, the campaign stuttered because of shortages of funds, trained personnel and vaccines. However, a renewed commitment to the programme in 1967 proved successful. The final patient with naturally acquired variola major smallpox was recorded in 1975, while the last to suffer variola minor smallpox from natural sources occurred in 1977. Happily, neither patient died. Today, the virus is confined to two high-security laboratories, one in the USA and the other in Russia.

Influenza

Numerous influenza (*Figure 2.14*) pandemics have occurred. Perhaps the most famous is the Spanish flu pandemic of 1918–1919, although the name has little relevance, which killed up to 40 million people, more than the nine million fatalities of World War I (WW1). It was one of the most devastating epidemics in history and is comparable to the Black Death (see *Box 2.2*) in sheer numbers of people killed. In two years, a fifth of the world's population was infected. The virus had a mortality rate of 2.5%, compared with that of the previous influenza epidemics of less than 0.1%. Further, most deaths were of people aged 20–40 years, which is unusual given that influenza normally kills the elderly and the very young. Indeed, the death rate for 15- to 34-year-olds from the influenza and associated pneumonia was 20 times higher in 1918 than in previous years.

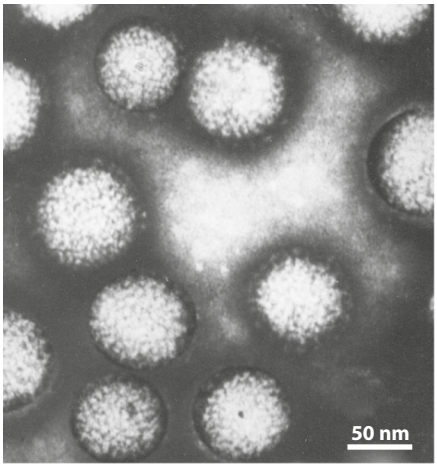


Figure 2.14 Electron micrograph of influenza viral particles. Courtesy of H. Cotterill, Manchester Royal Infirmary, UK. See also *Figure 2.10*. ⏎

The origin of this influenza variant is uncertain, but the mass movements of people resulting from WW1, such as the return of troops from Europe to their home countries along shipping lines, enabled the virus to spread rapidly. Outbreaks swept through North America, Europe, Asia, Africa, Brazil and the South Pacific. The shortage of medical facilities created by WW1 accentuated problems. As with the Black Death, social problems occurred, with shortages of coffins, morticians and gravediggers, so that bodies had to be stored in piles until they could be buried.

The genes of all influenza viruses are maintained in wild aquatic birds. Periodically, these viruses are transmitted to other species. Thus, the potential for outbreaks of influenza is still present. Influenza viruses mutate constantly but usually in such a way that the vaccine used in one year offers some protection against the next year's strain. However, every 10–20 years, a major mutation produces a new, particularly virulent strain against which the current vaccine offers little protection. Such viruses are associated with epidemics and pandemics like that of 1918–1919. Indeed, in 1997, epidemiologists and public health officials recognised a new variety of influenza virus, known as subtype H5N1, from its surface proteins. Given its lethal effects on poultry it was called chicken Ebola (see below) but is now more commonly called bird flu. When it infected the human population of Hong Kong it killed 6 of the first 18 confirmed cases. Fortunately, the H5N1 subtype cannot be transmitted through air from one human host to another. In Hong Kong, bird-to-human contact is relatively easy given the often close proximity of the two and is believed to have been the route of transmission. In 1997, Hong Kong officials ordered the

slaughter of its entire poultry population. However, by 2005/2006, cases of bird flu were being reported in diverse parts of the world. In 2023, a major outbreak in the UK killed many thousands of wild birds.

Viral epidemics and pandemics of the last 20 years

Viral epidemics that have become pandemics in the last two decades include:

- severe acute respiratory syndrome (SARS)
- influenza A H1N1
- Middle East respiratory syndrome (MERS)
- Ebola haemorrhagic fever/disease
- coronavirus disease 2019 (COVID-19)

Severe acute respiratory syndrome is caused by the virus SARS CoV-1, a member of the *Coronaviridae*. The first case of this respiratory infection was detected in Guangdong, China, in 2002 but spread to cause the 2002–2004 SARS pandemic affecting 30 countries. The virus was reported to have originated from bats found in the Yunnan province of China. Human-to-human transmission is via respiratory droplets. Symptoms of SARS include fever, dry cough and shortness of breath and, indeed, affected patients can develop pneumonia, which can be fatal. The pandemic involved 8000 patients, of which 700 died. It caused widespread social upheaval and had a considerable economic impact. Affected patients were treated to reduce and control symptoms, and its spread was reduced by using preventative measures, such as isolating confirmed cases and wearing protective clothing. Fortunately, there have been no reported cases of SARS since 2004, since there is currently neither an effective antiviral therapy nor a vaccine against the virus.

The influenza A H1N1 virus causes acute respiratory distress syndrome (ARDS), although the condition is commonly called swine flu because it resembles a respiratory infection found in pigs. The virus was the cause of a pandemic in 2009 that originated in Mexico and went on to affect more than 214 countries over an 18-month period before ending by 2010. Swine flu is a lung infection with symptoms including aching muscles, coughs, chills and fever, headaches, sore throat, watery eyes and diarrhoea. Antiviral therapy and an effective vaccine are available to treat and control the spread of swine flu; however, it still caused the death of more than 18 000 people in 2012, according to the WHO.

Middle East respiratory syndrome is caused by the coronavirus MERS-CoV (*Figure 2.15*). This virus spread from camels to humans, so the condition is sometimes referred to as camel flu. The first patient identified with MERS in 2012 was in Saudi Arabia and, indeed, most cases occurred there, although 26 other countries were affected by the pandemic, which resulted in nearly 900 deaths. Symptoms of MERS include coughing and difficulties breathing, fever, diarrhoea and vomiting. Transmission of the virus from one human to another is uncommon and only occurs during contact with severely ill patients. The pandemic was controlled by isolating affected patients and the use of local containment measures. Currently, antiviral drugs and vaccines are not available for MERS. The WHO has advised against eating uncooked or undercooked camel meat to prevent its spread.

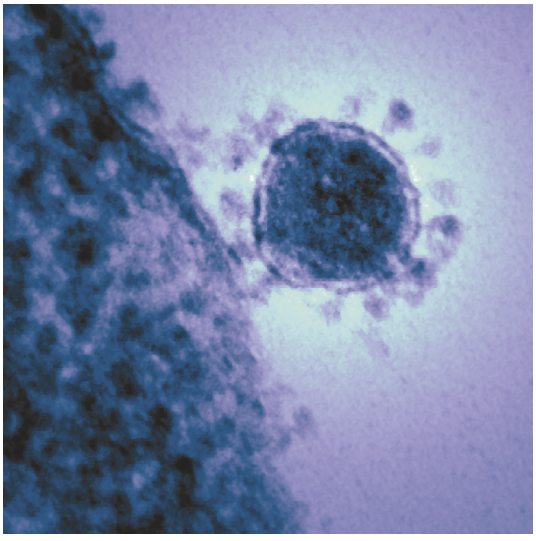


Figure 2.15 False colour electron micrograph of a Middle East respiratory syndrome (MERS) virion. Courtesy of the National Institute of Allergy and Infectious Diseases.

The Ebola virus causes Ebola haemorrhagic fever (EHF), which is also called Ebola virus disease (EVD). The pandemic associated with the Ebola virus was first reported in Guinea before spreading throughout western Africa and to the UK, USA and Italy during 2013–2016. The Ebola viruses responsible for this pandemic originated in infected bats and spread between humans via direct contact with blood or other body fluids from infected patients. Symptoms of EHF include abdominal pain, diarrhoea, skin rash, presence of faecal blood, yellowing of eyes and skin, and bleeding from ears, nose and mouth. Ebola haemorrhagic fever has significant mortality, and the pandemic was associated with over 11 000 deaths. It was controlled by using strict containment measures, contact tracing and surveillance in the community. People were encouraged to use protective measures to reduce the spread of the disease, such as avoiding contact with infected patients and strict hand hygiene. There is no antiviral treatment for EHF, but an effective vaccine is now available, having received regulatory approval in 2019.

The COVID-19 pandemic was caused by SARS-CoV-2 (*Figure 2.16*). This virus was first identified during an outbreak in Wuhan, China, in 2019. The virus quickly spread to other parts of Asia and then worldwide. In 2023, the WHO declared the pandemic over, but deaths still occur, adding to the total of nearly seven million. Symptoms of COVID-19 include fever, dry cough, fatigue and changes in sense of smell and taste. The most serious consequence of COVID-19 is acute respiratory disease with pneumonia. Severe illness is more likely in elderly patients and in those who are immunocompromised or have other underlying health issues such as diabetes mellitus. The virus spreads rapidly, mostly due to breathing in air contaminated by droplets and airborne particles carrying the virus. The risk of infection is highest for those close to infected people but can also occur over longer distances in closed environments, for example indoors. Infected patients can spread the virus for up to 10 days, even if they themselves do not have symptoms. The virus readily undergoes mutations, which are being constantly monitored, as new variants possess varying degrees of virulence and infectivity. Measures introduced to limit the spread included regular hand hygiene, strict disinfection procedures, the wearing of personal protective equipment and avoiding contact with affected individuals. The pandemic was so serious and widespread that many countries introduced social distancing, travel restrictions and even the lockdown of communities. Currently, there is no effective cure for COVID-19 but therapies are available to reduce the severity of its symptoms. The successful introduction of vaccines together with the extensive immunity built up in surviving patients ('herd immunity') has had a huge impact on curtailing the spread of the disease and in reducing mortality.

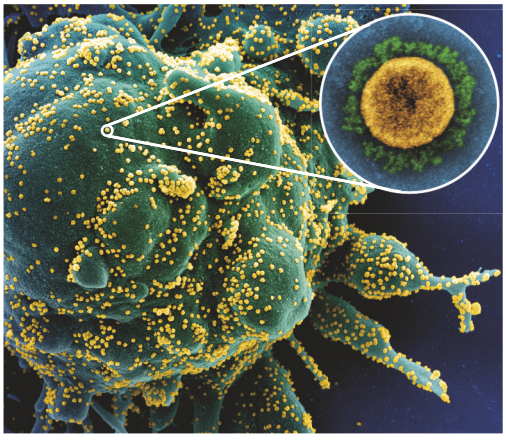


Figure 2.16 False colour electron micrograph of an apoptotic cell (see *Chapter 17*) infected with severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) virions, the causative agent of coronavirus disease 2019 (COVID-19), isolated from a patient sample. The insert shows a single virion. Image compiled courtesy of photographs from the National Institute of Allergy and Infectious Diseases.

2.7 BACTERIA

Bacteria (*Figure 2.17*) are responsible for a large proportion of all infectious diseases and, although prokaryotes share considerable biochemistry with eukaryotes, their metabolism differs in numerous ways, making them susceptible to chemical agents that do not significantly damage human cells. This is the basis of antibiotic therapy (*Chapter 3*). While most species of pathogenic bacteria are extracellular pathogens and do not enter host cells, some significant groups are intracellular parasites. Examples of the latter include mycobacteria and *Listeria*.

Evolution has provided pathogens and parasites with a wide range of factors that allow them to invade and colonise their host while at the same time avoiding and/or neutralising host defences.

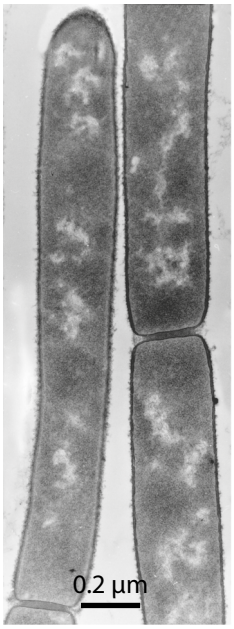


Figure 2.17 Electron micrograph of bacterial cells. Note the lighter stained nuclear material free in the cytoplasm and not contained in a nucleus. Courtesy of Dr I.D.J. Burdett, National Institute for Medical Research, London. ⏎

Microbial virulence factors

Many virulence factors of pathogens have been identified; some are relatively non-specific in action. For example, some microorganisms possess specialised iron uptake systems. Microorganisms require iron for oxygen transport, mitochondrial energy metabolism, electron transport, the biosynthesis of nucleic acids and gene expression. However, although an essential element, iron is often only available in limited quantities, and microorganisms that possess a variety of iron uptake systems are able to grow in regions of the host that would otherwise be expected to be sterile, since little iron is available. Other virulence factors have rather more defined defensive or offensive actions.

BOX 2.2 PLAGUE ⏎

The term **plague** was originally applied to any disease causing great mortality (see *Box 2.1*). However, it is now (almost) exclusively applied to the disease caused by the bacterium *Yersinia pestis* (*Figure 2.18*). *Y. pestis* is thought to have originated in Asia, diverging from the enteric bacterium *Y. pseudotuberculosis* 5000–6000 years ago. During its evolution, it acquired several potent pathogenic features. These included plasmids encoding formidable virulence-associated features, enzymes that disrupt responses of the immune system and an effective Fe-capturing system (*Section 2.6* and *Chapter 4*). The bacterium is mainly a pathogen of rodents and lagomorphs but, in humans, causes the potentially fatal bubonic and pneumonic plagues. For many years it was widely believed that the plague vector was solely rodent fleas carried on black rats. However, modern epidemiological modelling of the rapid spread of the disease during epi- and pandemics and vector–pathogen studies indicate that human fleas and body lice (*Figure 2.19*) are also culpable vectors.

Despite the name, body lice live in clothing next to the skin. They only visit the skin to feed, like fleas, on blood using specialised piercing mouthparts. Infection usually results from a bite from a human flea or body louse carrying the plague bacterium when they feed on blood or by handling an infected animal. The bacterium leaves the blood and colonises lymph nodes. Here, growth of the bacterium leads to fever and causes the nodes to painfully blacken and swell grotesquely, forming the characteristic buboes from which the disease derives its name. Rupture of the buboes releases bacteria back into the blood from where they can attack a variety of organs. The coughing and sneezing of infected individuals with damaged lungs spreads pneumonic plague; a much more lethal form

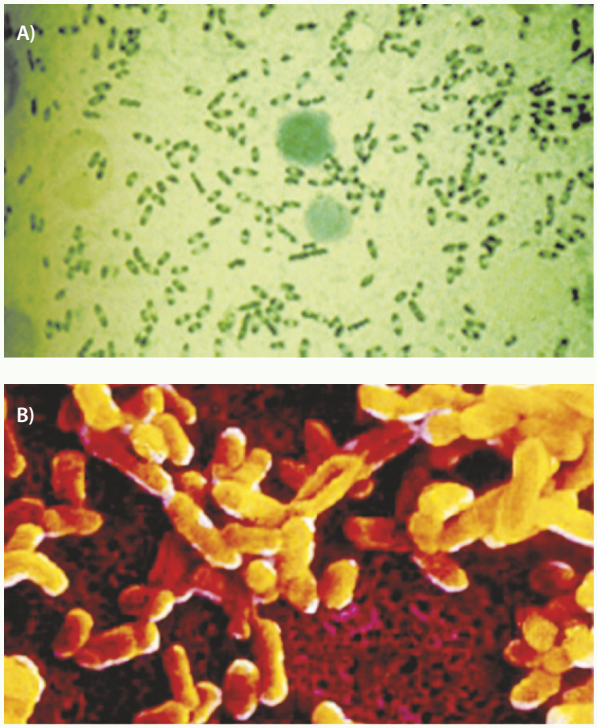


Figure 2.18 (A) Photo- and (B) electron micrographs of *Yersinia pestis*. The cells are 1–3 µm long. (A), Courtesy of Dr M.K. Khalid, Editor, *Middle East Journal of Emergency Medicine*; (B), courtesy of Dr M. Schneider, *Kryptozoologie*. ⏎

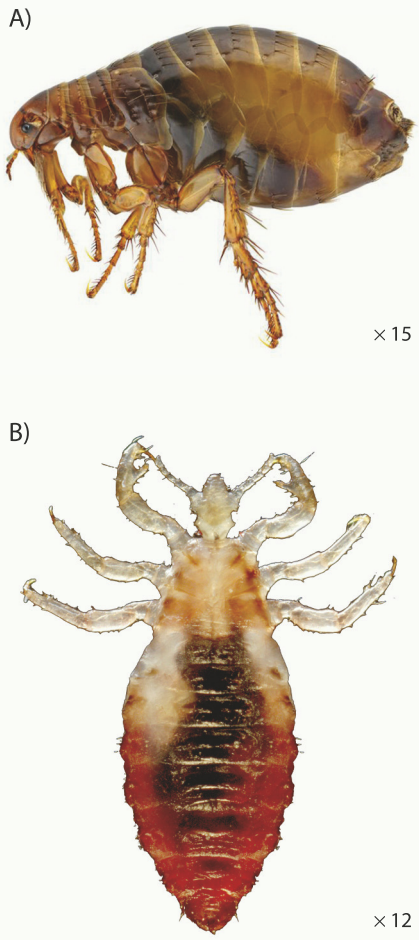


Figure 2.19 Photographs of a (A) human flea (*Pulex irritans*) and (B) human body louse (*Pediculus humanus*). (A), Courtesy of Shutterstock.com/Cosmin Manci; (B), courtesy of the Centers for Disease Control and Prevention/Dr F. Collins. ⏎

of the disease. Bubonic plague is fatal in over 30% of cases. In contrast, pneumonic plague is often fatal even with antibiotic therapy.

Bubonic plague has had a profound impact on humans throughout recorded history. Indeed, the three great outbreaks associated with *Y. pestis* are probably the worst biological disasters to have attacked humankind. The first epidemic, called the Plague of Justinian, originated in Egypt in 541–544 CE and continued sporadically until about 750 CE. It is thought to have killed 50–60% of the populations of the eastern Roman and the Sasanian empires.

The second great plague began in China in the 1330s. At the time, China, like now, was one of the world's busiest trading nations. The disease gradually spread to western Asia and the Middle East and entered western Europe in October 1347 on Italian merchant ships returning from Caffa on the Black Sea coast of Crimea to Messina, which was a key link in trade routes with China. Within days, the disease spread to the city and surrounding countryside. In August the following year, the plague had spread as far north as England, where people called it the **Black Death** because of the black buboes on patients. Medieval medicine was ineffective in treating the disease and extensive prayers and sacrifices to God/gods did not elicit a response. However, in winter, the disease declined simply because the insects that carried the bacteria were dormant. Each spring, the plague returned. After five years, more than 13 million people in China and 20–30 million in Europe (approximately 40% of the European population) were dead. In addition to these terrible fatalities, the Black Death produced enormous social changes from which the structure of medieval society never recovered. The resulting serious labour shortages all over Europe led to demands for higher wages by workers, and peasant revolts broke out in Belgium, England, France and Italy. The end of the 1300s saw the eventual collapse of the prevailing social system of tied serfs. Even when the Black Death epidemic ended, smaller outbreaks of bubonic plague continued until the eighteenth century. For example, in England, an outbreak in 1665–1666 is estimated to have killed 200 000 people out of a population of roughly 4 000 000.

In 1855, the third major plague pandemic started in China, which initiated a series of outbreaks lasting until the twentieth century. In 1894, outbreaks in Guang Zhou (then called Canton) and Hong Kong killed 80 000–100 000 people, and within 20 years the plague had spread and killed over 10 million people worldwide.

Bubonic plague has an incubation time of approximately seven days and, if not treated promptly, bacteraemia develops rapidly, with mortalities of 30–75%. Pneumonic fever is highly infectious and is spread by aerosols and droplets when diseased patients cough and sneeze. It has a shorter incubation of two to three days and a mortality of approximately 90% when untreated. Outbreaks of plague still occur in some rural communities in underdeveloped countries. The World Health Organization (WHO) reports 1000–3000 cases of plague annually. It is unlikely that the disease will ever be completely eradicated because wild animals hold a huge reservoir of the bacterium. However, advances in living conditions, public health and antibiotic therapy would seem to make future pandemics of *Y. pestis* unlikely. For example, effective laundering of clothes has made the human body louse an endangered species.

Defensive virulence factors

Numerous pathogens evade host defences by producing slime layers or possessing polysaccharide capsules (*Figure 2.20*). Slime layers consist of exopolysaccharides (EPSs) that bind large quantities of water. Slime production is particularly important in bacteria that form biofilms, since it forms a protective coat around the bacterial population. For example, the biofilm formed by the opportunistic *P. aeruginosa* in the respiratory tract of cystic fibrosis patients protects it from the immune system and antibiotics. Capsules

generally consist of a single type of polysaccharide that binds considerable quantities of water, forming a layer around the bacterial cell that protects it from molecules of the immune system. The polysaccharide is often negatively charged, which renders it resistant to uptake by phagocytic cells (*Chapter 4*). Capsules can also protect the bacterium from the immune system (*Chapter 4*). Some polysaccharide capsules are molecular mimics of host cell surface structures. The capsule of *E. coli* K1 and the type B capsule of *Neisseria meningitidis* consists of α,2→8 *N*-acetylneuraminate residues that are identical to those on the neuronal adhesion molecule N-CAM and other sialylated molecules of the nervous system. Consequently, host immune systems do not recognise the bacteria as foreign, and both pathogens can invade the central nervous system (CNS) causing meningitis.

Other examples of molecular mimicry are the many proteins of pathogenic bacteria that are homologous to specific regions of host proteins. *Y. pestis* induces the production of antibodies that cross-react with part of a particular variant of a host protein called human leukocyte antigen (HLA)-B27 (*Chapters 4* and *5*). Cross-reactivity between other bacterial species and HLA-B27 is thought to be involved in the development of types of arthritis known as Reiter syndrome and in ankylosing spondylitis.

Some bacteria directly or indirectly activate or suppress actions of the immune system by producing pathogenicity factors called modulins or microkines. The P fimbriae of the uropathogenic *E. coli*, for example, induces an increase in the release of interleukin-4 by uroepithelial cells. Modulation of cytokine production may lead to increased pathogenicity (*Chapter 4*).

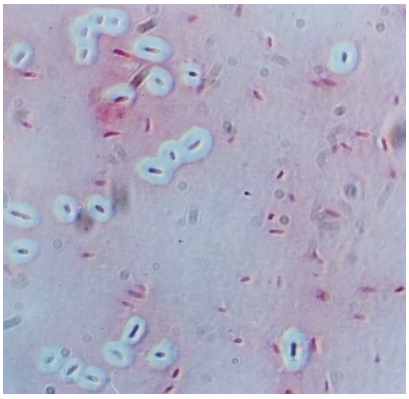


Figure 2.20 Photomicrograph of *Bacteroides fragilis,* an obligate anaerobic bacterium that is normally found in the gastrointestinal tract. This bacterium is most frequently isolated from clinical infections such as peritonitis (Chapter 11). Note the prominent capsules surrounding each cell. Courtesy of Dr S. Patrick, Queen's University Belfast, UK.

BOX 2.3 EVADING IMMUNE SYSTEMS

A number of viruses and microbial pathogens/parasites prevent their hosts from mounting an effective immune response by changing their surface antigens. This can occur in a number of ways. For example, several viruses, including influenza, HIV and SARS-CoV-2, have genes coding for surface proteins that have hypermutability, meaning they mutate at relatively fast rates. Thus, the antigenic structure of these surface proteins is prone to change at intervals, leaving a host population that is less or no longer immune to that virus. Other microorganisms that undergo antigenic variation include the trypanosome *Trypanosoma brucei*, which causes sleeping sickness. These regularly change the structure of their surface glycoproteins during the course of an infection. The malarial parasites *Plasmodium* also express different surface antigens during the infection process. Some parasitic worms, for example schistosomes, become coated with patient antigens, such as major histocompatibility complex (MHC; also often referred to as the HLA system) molecules and common blood group antigens (*Chapter 6*), and so avoid recognition by the host.

Offensive virulence factors

Bacterial offensive virulence factors include:

- adhesins
- invasins
- toxins

Adhesins are proteins found on the surfaces of microbial cells that bind to specific sites on the cells of the host. The best studied are those in the pili of, for example, certain strains of *E. coli*, *Vibrio cholerae* and *Helicobacter pylori*. Pili are fibres about 2 μm long and 2–8 nm in diameter that extend from some bacterial cells (*Figure 2.21*). They consist of about 1000 protein molecules, which include adhesin glycoproteins called **lectins**, which can bind specifically to certain sugars or glycosidic bonds found in some carbohydrates. The lectins in pili *E. coli* and *V. cholerae* bind to mannose and fucose sugar residues, respectively, that are found on the surfaces of host cells.

Invasins are proteins that allow bacterial pathogens bound to the host to be internalised, that is, enter the host cell. This prevents the bacterial cell being removed by ciliary action or washing and ensures that the pathogen is protected from direct immune attack. Once internalised, the microorganisms may remain in membrane-bound vesicles. Others escape into the cytosol and so avoid the killing mechanisms associated with phagocytosis. Some microorganisms are so adapted to intracellular life that they are unable to reproduce outside their host cell. These include species of *Chlamydia* and

Figure 2.21 Electron micrograph of an *Escherichia coli* cell with numerous pili.

Rickettsia and some mycobacterial pathogens, which are therefore obligate intracellular parasites. Some of these types of organisms cause infectious diseases, for example *Chlamydia pneumoniae*, *Rickettsia typhi* and *Mycobacterium leprae*. However, others, such as strains of Listeriae, Salmonellae, Shigellae and Yersiniae, are facultative and can live outside their host cells.

Toxins produced by different types of pathogenic microorganisms can be classified into two types:

1. cell-associated toxins, for example endotoxins
2. those secreted by the bacterium, called exotoxins

Endotoxins are produced by Gram-negative bacteria. They are a toxic portion of the lipopolysaccharides embedded in the outer membrane of their cell walls called lipid A (*Figure 2.22*). The clinical damage caused by lipid A varies with the susceptibility of the host. Fever is common because endotoxins stimulate host cells to release cytokines that affect the thermoregulatory centre of the hypothalamus. In serious cases, endotoxic shock can result.

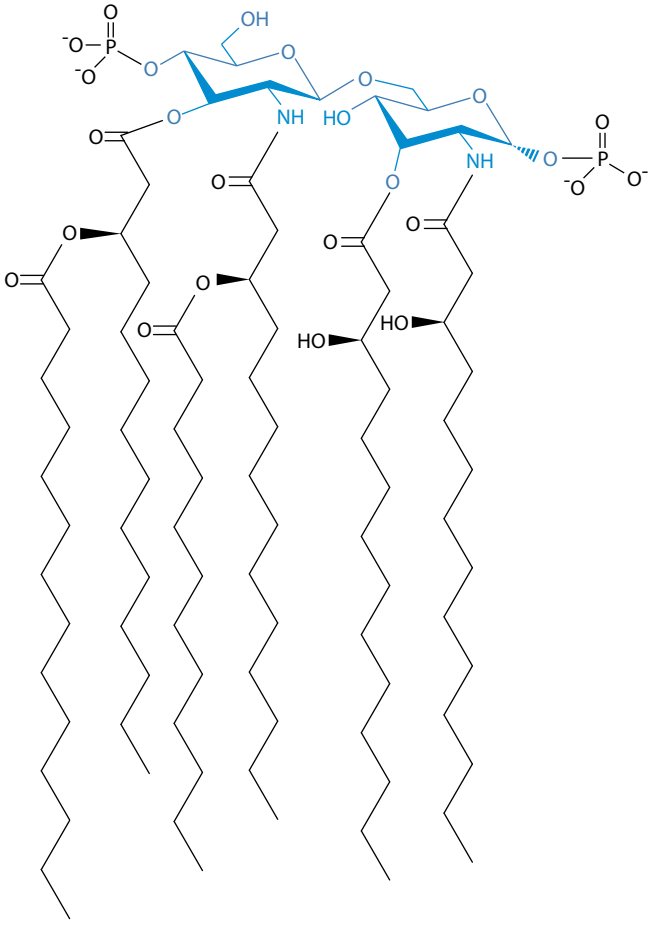


Figure 2.22 Structure of the *Escherichia coli* lipid A virulence factor.

Exotoxins are proteins secreted by both Gram-positive and -negative bacteria. They may be subdivided into three groups:

1. those that damage membranes
2. those with specific host targets
3. superantigens

Exotoxins that damage membranes cause the cell to gain or lose ions and water. In high doses, cell lysis occurs, hence they are sometimes called haemolysins or cytolysins. A clinically significant feature of such toxins in an infection is their antiphagocytic activity. Some of the haemolysins and cytolysins of bacteria, such as *Staphylococcus aureus*, *Streptococcus pyogenes* and *Bordetella pertussis*, are polypeptides that aggregate in the

membranes of host cells forming pores (*Figure 2.23*). Thiol-activated lysins are predominantly produced by Gram-positive bacteria. These toxins are proteins containing large numbers of cysteine residues. They bind to cholesterol molecules in membranes of target cells and form aqueous pores comprising 25–100 toxin molecules that constitute damaging lesions in the host cell.

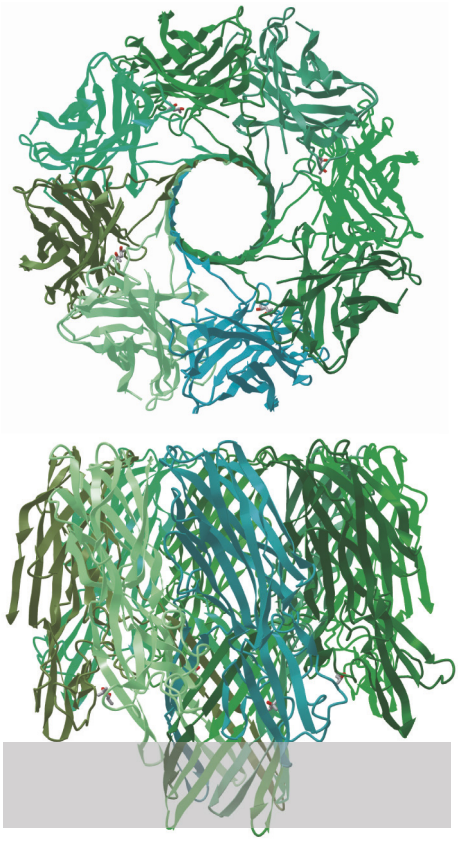
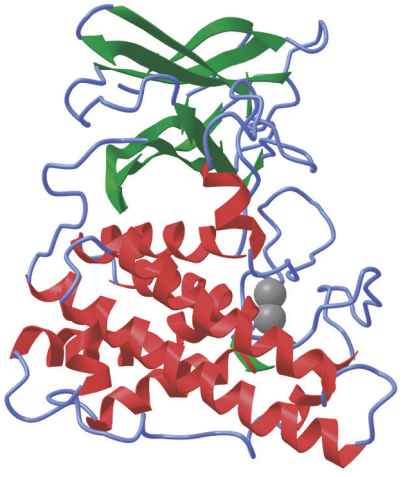


Figure 2.23 Molecular model of the membrane pore formed by an octamer of haemolysins from *Staphylococcus aureus*. The approximate position of the lipid bilayer is shown in grey. PDB file 3B07.

Phospholipases of pathogens catalyse the hydrolysis of phospholipids in the membranes of host cells. The α toxin of *Clostridium perfringens*, for example, is a phospholipase that catalyses the hydrolysis of phosphatidylcholines (*Figure 2.24*). Similarly, the β haemolysin of *S. aureus* is a sphingomyelinase that promotes the breakdown of sphingomyelins (*Figure 2.25*). The degradation of membrane lipids, naturally, results in a loss of membrane integrity and function.

A)



B)

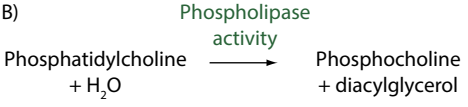


Figure 2.24 (A) Molecular model of the α toxin of *Clostridium perfringens*. The two grey spheres represent Zn^{2+}. PDB file 1QM6. (B) An example of the reaction catalysed by phospholipases, such as the α toxin of *Clostridium perfringens*.

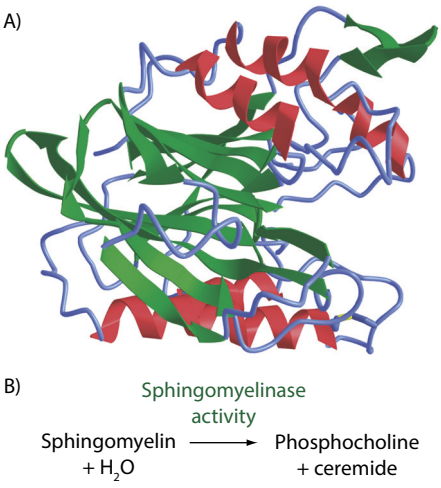


Figure 2.25 (A) Molecular model of the β haemolysin of *Staphylococcus aureus*. PDB file 3I5V. (B) An example of the reaction catalysed by a sphingomyelinase, such as the β haemolysin of *Staphylococcus aureus*.

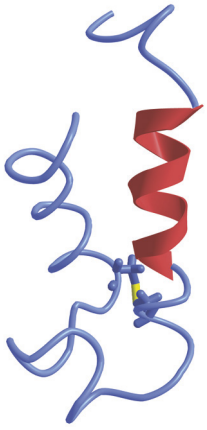


Figure 2.26 Molecular model of the heat-stable enterotoxin of *Escherichia coli*. PDB file 1EHS.

Exotoxins that target specific sites in the host do so in a wide variety of ways; for example, they may act to deregulate or kill cells or they may have an extracellular target. Gram-negative bacteria, such as Enterobacteriaceae (*E. coli*, *Citrobacter freundii*, *Yersinia enterocolitica*), secrete heat-stable enterotoxins as small as M_r 2000 (*Figure 2.26*). These toxins bind to specific receptors that are part of a cyclic GMP-dependent signal transduction system of enterocytes in the upper intestinal epithelium. This system regulates the concentration of intracellular cyclic GMP that, in turn, is involved in the activation of intracellular enzymes, for example protein kinase G (*Chapter 7*). The binding of the toxins interrupts the secretion of Na^+ and Cl^- and this results in a watery diarrhoea.

Clostridium botulinum and *tetani* toxins consist of two dissimilar polypeptides usually referred to as its A and B subunits. The B subunit recognises and binds to specific target cells and facilitates entry of the A subunit into host cells. The A subunits are Zn-dependent proteases and are neurotoxic. Botulism is a potentially fatal type of food poisoning. It results from the effects of seven related AB-type toxins (*Figure 2.27*), which are regarded as the most lethal biological toxins: it is estimated that 1 mg is sufficient to kill 1 000 000 guinea pigs. Botulinum toxin is absorbed by the GIT and is transported to presynaptic membranes of motor neurons, where it inhibits the release of acetylcholine, causing a flaccid paralysis and death by suffocation. Improvements in care have reduced mortality rates for botulinum poisoning to 5–10%; however, prolonged treatment is required. It consists of assisted ventilatory support and antibiotic therapy (*Chapter 3*) and botulinum antitoxin to neutralise free toxin in the circulation.

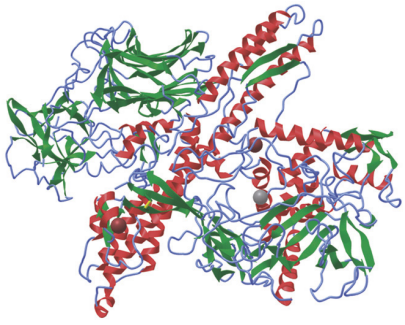


Figure 2.27 Molecular model of a botulinum toxin. The bound Ca^{2+} and Zn^{2+} are shown as red and grey spheres, respectively. PDB file 1S0F.

In contrast with botulism, tetanus toxin (tetanospasmin [*Figure 2.28*]) enters patients typically through deep cuts. Following its internalisation into neurons of the CNS, tetanospasmin migrates to the peripheral endings of inhibitory nerve fibres by retrograde axonal transport. Here it prevents the release of the inhibitory neurotransmitters γ aminobutyrate and glycine, which oppose the action of acetylcholine. This leads to unchecked excitatory impulses and a continuous stimulation of muscles, producing the painful spastic paralysis typical of tetanus. Continued contractions of mouth muscles restrict jaw movements producing lockjaw, a colloquial name for the condition. In 30–50% of cases, prolonged contractions of respiratory muscles cause death by asphyxiation. Symptoms must be clinically managed because the toxin is protected when bound to nerve fibre endings. Treatment consists of antibiotics and anti-tetanus immunoglobulin to neutralise any unbound toxin. An effective prophylaxes vaccination regimen is available (*Chapter 3*).

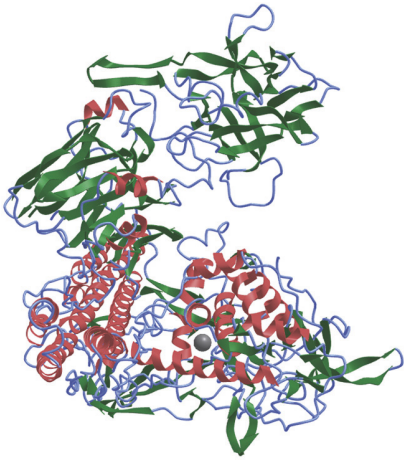


Figure 2.28 Molecular model of the tetanus toxin, tetanospasmin. The bound Zn^{2+} is shown as a grey sphere. PDB file 5N0C.

The cholera, diphtheria and pertussis exotoxins are ADP ribosyltransferases that use NAD^+ as a donor substrate. The ADP ribosyl portion is transferred to the target protein, releasing nicotinamide. Cholera and pertussis toxins attack G proteins so that G protein-mediated signal transduction pathways become dysfunctional. Thus, the cholera toxin (choleragen; *Figure 2.29*) catalyses the ADP ribosylation of arginine residues of the G_α subunit of the adenylate cyclase transduction system (see *Figure 7.11*) fixing it in its active form. This leads, in turn, to a long-lasting activation of adenylate cyclase and biosynthesis of cyclic AMP, and activation of protein kinase A. The net result is a long-lived opening of the chloride channel of the cystic fibrosis transmembrane conductance regulator (*Chapter 16*), which increases secretion of hydrogen carbonate (HCO_3^-) and Cl^- into the intestinal lumen but which inhibits the absorption of Na^+ and Cl^-. The resulting osmotic effect causes a massive leakage of intracellular water into the intestinal lumen

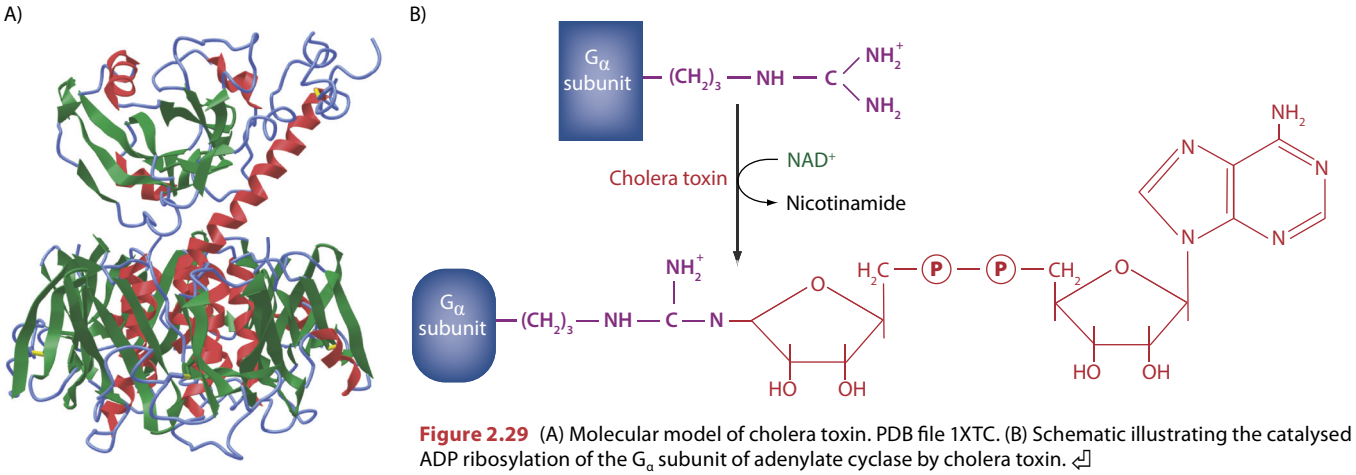


Figure 2.29 (A) Molecular model of cholera toxin. PDB file 1XTC. (B) Schematic illustrating the catalysed ADP ribosylation of the G_α subunit of adenylate cyclase by cholera toxin.

and a subsequent fulminant diarrhoea, resulting in acid–base disorder, hypokalaemia and hypovolaemia (*Chapters 8* and *9*). The mortality can be as high as 70% from heart and kidney dysfunction in untreated cases. However, if patients receive prompt fluid and electrolyte replacement therapy and suitable antibiotic treatment, death rates are reduced to less than 1%. Improvements in public health, particularly clean water supplies, have drastically reduced the incidence of cholera in the developed world.

Corynebacterium diphtheriae causes diphtheria. Its toxin targets a single protein: the elongation factor EF2, which functions in translating mRNAs. It is thought that a single toxin molecule can disable every ribosome in a susceptible cell by its ADP ribosylation activities (*Figure 2.30*) leading to cell death. Diphtheria was once a common, potentially fatal, disease of childhood. However, the development of an effective vaccine and its efficient dissemination in early childhood (*Chapter 3*) has virtually eliminated the problem. Cases in the UK are now the result of foreign travel to less developed areas.

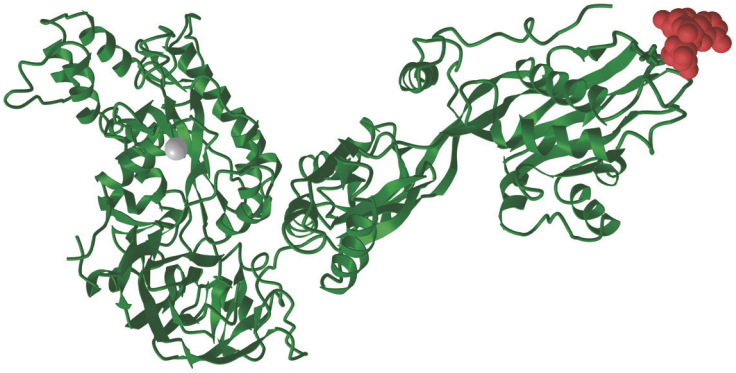


Figure 2.30 Molecular model of an ADP ribosylated elongation factor 2. The bound adenosine 5′-diphosphoribose group is highlighted in red. The bound Mg^{2+} is represented as a grey sphere. PDB file 1U2R.

The Shiga family of toxins produced by *S. dysenteriae* and *E. coli* (*Figure 2.31*), like the diphtheria toxin, also inhibit protein synthesis but by a different mechanism. They are *N* glycosidases that hydrolyse an *N*-glycosidic bond between specific adenine and ribose residues of the 28S ribosomal RNA of the 60S ribosomal subunit (see *Figure 3.47*).

Several microbial toxins catabolise components of connective tissues. Staphylococci and Streptococci secrete enzymes that degrade the extracellular matrix, allowing them to invade and colonise tissues. *C. perfringens* releases a variety of carbohydrases and proteases that also degrade connective tissues, allowing it to colonise and spread through subcutaneous tissues. If the bacterium invades deeper muscle tissue it may cause necrosis and release gas from anaerobic fermentations (gas gangrene). This type of infection is extremely painful and can spread rapidly. Much of the direct damage is caused by the α toxin described above. Other pathogens secrete proteases that degrade specific components of the immune defence, such as immunoglobulin A (IgA) molecules, or catalases and superoxidase dismutases (SODs) that are used by some immune cells to kill microbial pathogens (*Chapters 4* and *13*).

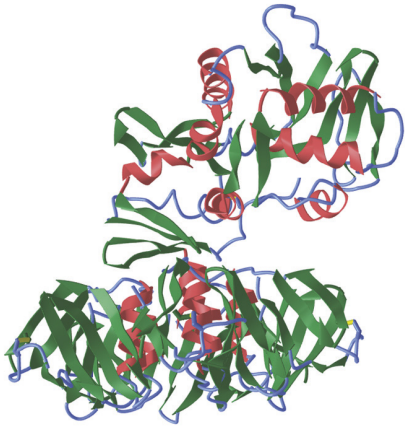


Figure 2.31 Molecular model of Shiga toxin produced by *Shigella dysenteriae*, the pathogen responsible for severe types of dysentery. PDB file 1DM0.

Superantigens are proteins secreted by Gram-positive pathogens, such as some strains of *S. aureus* and *S. pyogenes*. They are extremely effective and potent stimulators of the immune system because of their unique ability to stimulate large numbers of its cells simultaneously. This leads to a massive release of cytokines that activate numerous physiological systems, such as the temperature regulatory system (see *Box 4.6*). The superantigen of *S. pyogenes* is responsible for streptococcal toxic shock syndrome (STSS). The fever, shock and tissue damage associated with STSS is thought to be the result of an overproduction of the cytokines tumour necrosis factor-α, interleukin-1β and interleukin-6. The superantigen of *S. aureus*, the pyrogenic toxic shock syndrome toxin 1 (*Figure 2.32*), causes symptoms that can lead to a rapid failure of many body organs and death.

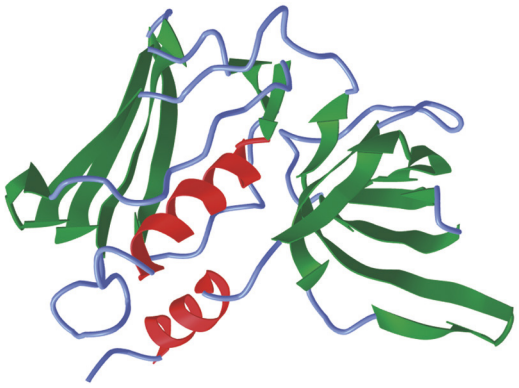


Figure 2.32 Molecular model of toxic shock syndrome toxin 1 (TSST-1) produced by *Staphylococcus aureus*. PDB file 2QIL. ⏎

2.8 PROTOZOA

Protozoa are single-celled eukaryotes, although, in some species, these may group together as loose aggregates. Malaria, caused by species of *Plasmodium*, is easily the most important clinical protozoal infection worldwide and is responsible for 100 000s of deaths annually (*Box 2.4*).

BOX 2.4 MALARIA: THE BAD AIR DISEASE ⏎

Malaria infects erythrocytes and hepatocytes. It is named malaria from the eighteenth century Italian *mala aria*, meaning 'bad air', from the then belief it was caused by the unwholesome air of swampy districts. In tropical climates particularly, malaria causes over 400 000 deaths and 200 million new cases annually. Malaria is caused by four main *Plasmodium* species: *P. falciparum*, *P. malariae*, *P. ovale* and *P. vivax*. African great apes incubate a number *Plasmodium* spp., to which some of the human parasites are related. For example, *P. falciparum* has been shown to have evolved from a gorilla parasite no more than 10 000 years ago.

The life cycle of *Plasmodium* (*Figure 2.33*) is complicated by any standards! A feeding female *Anopheles* mosquito infects a human with a haploid stage of the parasite called a sporozoite. These infect liver cells, replicate and develop, and when released as merozoites can infect other liver cells or erythrocytes. Here the parasite, usually called a trophozoite, also grows and divides, leading to the eventual lysis of the erythrocyte. The released trophozoites can invade other erythrocytes in 48-hour cycles of invasions and lyses that cause the characteristic fevers and chills associated with malaria. However, eventually, some trophozoites develop into male and female gametocytes that can be ingested by a feeding mosquito. Within the insect gut, the gametocytes fuse, forming a diploid zygote. The zygote enters the gut wall, forming a cyst. Within the cyst, the zygote develops and forms sporozoites that migrate to the salivary gland. Thus, the cycle of events can be repeated when the mosquito next feeds on a human.

Symptoms of malaria may present after a variable incubation period and include headaches, general malaise, sweating, muscular pains and rigours, and anorexia, and may result in death. The clinical course and symptoms of different forms of malaria are variable, although infections are characterised by recurrent attacks of chills and fevers, which are usually most severe during the lysis of erythrocytes.

Prophylactic drugs and especially the use of insecticide-impregnated mosquito nets offer some control over new malaria cases. Also, a variety of antimalarial drugs is available (*Chapter 3*). Some prevent the hepatic forms of *Plasmodium* from invading erythrocytes; others destroy the erythrocytic or gametocyte forms of the parasite in patient blood, preventing transmission of the parasite by the mosquito. Chloroquine then sulphadoxine–pyrimethamine were the usual choices for treating malaria because of their low cost, safety and initial effectiveness. Drugs to treat those parasites that developed resistance to them included halofantrine, mefloquine, quinine, quinidine and others. However, strains of *P. falciparum* resistant to these single drug treatments are now endemic. During 2010–2015, deaths due to malaria decreased by 20–30% following the introduction of combination therapies using artemisinin or its derivatives artemether, artesunate or dihydroartemisinin, combined with amodiaquine, lumefantrine, mefloquine, sulphadoxine–pyrimethamine or piperaquine (artemisinin-based combination therapies [ACTs]). Artemisinin is isolated from the Chinese medicinal herb *Artemisia annua* and rapidly eliminates susceptible

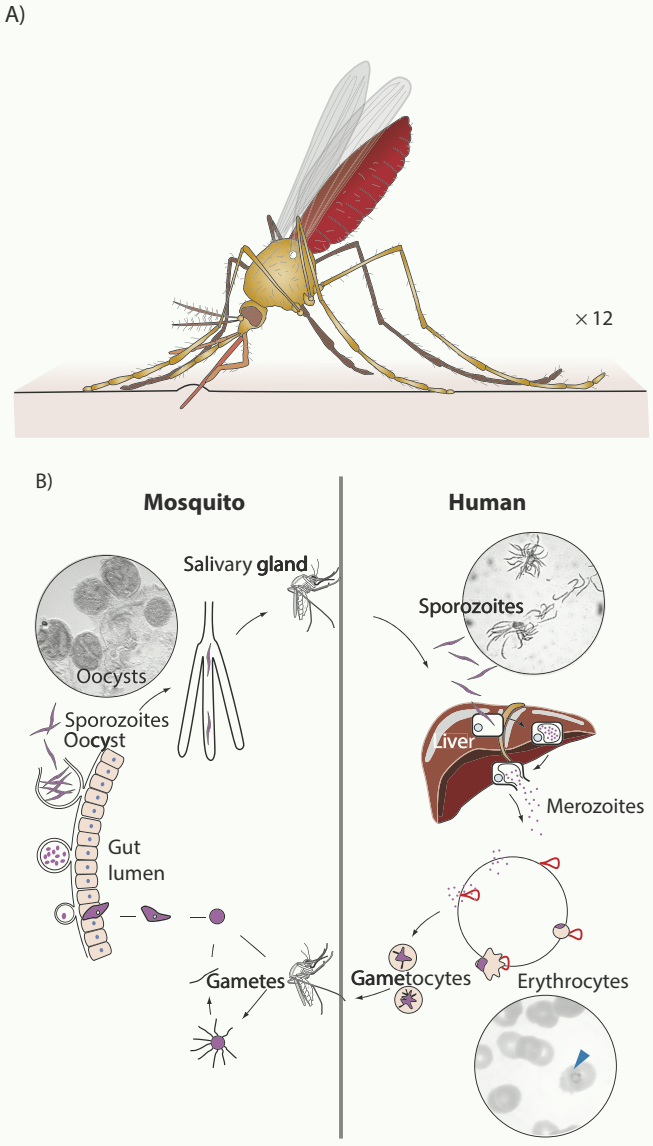


Figure 2.33 (A) An *Anopheles* mosquito. (B) Life cycle of a malarial parasite. The three inserts from top left in a clockwise direction are oocysts in the *Anopheles* gut, free sporozoites in the blood of a human patient and a young *P. falciparum* trophozoite in an erythrocyte. Note its characteristic signet ring shape. Courtesy of J.R. O'Kecha, Homerton University Hospital, London.

parasites from the blood. The second group of antimalarials have a longer lasting antimalarial benefit (*Figure 2.34*). The rationale for combination therapy is that resistance will be less likely to develop if each component operates by a different mechanism (*Chapter 3*). In many cases, neither the molecular mechanisms of action of antimalarial drugs nor the biochemical mechanisms of resistance to them were well understood. However, recent years have seen advances in both aspects, which should be of benefit in developing new combinations of fast-acting drugs to replace artemisinins and ones with different, slower-acting antimalarial properties to partner them.

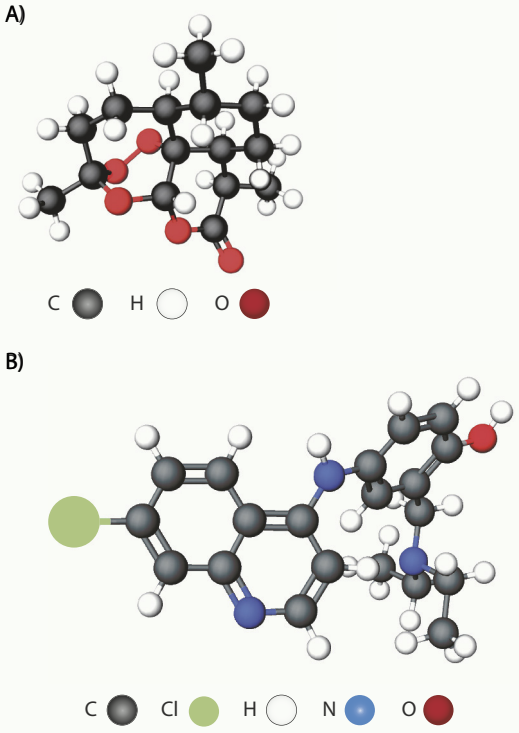


Figure 2.34 Computer drawn structures of the antimalarial drugs (A) artemisinin and (B) amodiaquine.

The remarkably complicated life cycle of the malarial parasites, with their prehepatic, hepatic, pre-erythrocyte and erythrocyte stages, means there is a large choice of antigenic targets to use for vaccine development (*Chapter 3*). Despite several efforts and trials to develop an effective malarial vaccine, success has generally been limited. However, complete sequencing of the parasite genome in 2005 was a cause for optimism. Indeed, following extensive trials in Ghana, Kenya and Malawi, in October 2021 the World Health Organization (WHO) recommended the use of a new malaria vaccine called RTS,S/AS01 in children living in regions with moderate to high incidences of *P. falciparum* malaria. Four doses of RTS,S have been shown to reduce cases of clinical malaria by 39% and severe malaria by 30%. A second new vaccine called R21 devised by the University of Oxford appears to be even more effective. A preliminary trial involving about 400 children in Burkina Faso indicated it was 80% effective when administered initially in three doses with a booster a year later. This was followed by a much larger trial with approximately 5000 children. The results of this second trial have been released to some African countries and are thought to confirm the findings of the earlier studies. Following its assessment of these data on the safety and effectiveness of R21, Ghana became the first country to approve its use in children 5–36 months old. Other African countries and the WHO are currently also considering approving its use. The promising development of these new vaccines could greatly contribute to the WHO global technical strategy of reducing the incidence of malaria by at least 90% by 2030.

Protozoa are, however, the causative agents of other infections including:

- trypanosomiases
- amoebiasis
- giardiasis
- leishmaniasis
- toxoplasmosis
- trichomoniasis

Trypanosomiases

Trypanosomiases are African trypanosomiasis, commonly called sleeping sickness, and Chagas disease, also referred to as American trypanosomiasis (see *Box 2.5*). Both are caused by parasitic trypanosomes using insects as vectors. African trypanosomiasis is caused by two subspecies of *Trypanosoma brucei*: *T. brucei gambiense* in western and central Africa and *T. brucei rhodesiense* in eastern and southern African regions (*Figure 2.35*). The disease vector is the tsetse fly (*Glossina* spp.; see *Figure 2.45*), which is found only in sub-Saharan Africa. Chagas disease is caused by *Trypanosoma cruzi* and, in contrast, is not spread by flies but by *Triatoma* bugs (see *Box 2.5*).

The incidence of African trypanosomiasis appears to be decreasing; nearly 10 000 cases occurred in 2009 but, in 2020, the WHO reported fewer than 700. Like mosquitos, tsetse flies feed on blood using piercing mouthparts enabling an infected fly to inject metacyclic trypomastigote stages of the trypanosome into blood or skin tissues. Within the blood, the parasite transforms into a trypomastigote stage, which replicates by binary fission in body fluids, that is, blood, lymph and spinal fluid. Tsetse flies become infected with bloodstream trypomastigotes on taking a blood meal from an infected patient. The protozoan then differentiates into procyclic trypomastigotes in the midgut of the fly. These also reproduce by binary fission. Trypomastigotes leave the midgut and form yet another stage, called epimastigotes, that migrate to the salivary glands of the fly and continue to multiply by binary fission, transforming into metacyclic trypomastigotes. The cycle is completed when these are injected into a new host when the fly feeds. Thus, unlike malarial parasites, the life cycle of *T. brucei* is entirely extracellular.

Infection presents in two stages: first an initial haemolymphatic phase that is followed by a second meningoencephalitis stage after the trypanosome has invaded the CNS. However, it is difficult to distinguish between the two, as many of the signs and symptoms of both are similar. The first stage incubation of *T. brucei rhodesiense* is one to three weeks but that of *T. brucei gambiense* progresses more slowly, although the first stage symptoms for both include fatigue and weakness, headaches, a general malaise, pruritus and arthralgia, and can include hepatosplenomegaly, loss of weight and intermittent fevers lasting one to seven days. The intervals between fevers can be days or even months. The second stage of *T. brucei rhodesiense* infections starts after 21–60 days; *T. brucei gambiense* is slower at 300–500 days. Additional to the first stage symptoms is a reversal of the sleep–wake cycle, hence sleeping sickness, and is characterised by daytime somnolence, nocturnal insomnia and sudden urges to sleep, which is, unsurprisingly, accompanied by anxiety, attention deficit, hallucinations and delirium that can progress to coma. The condition is fatal if untreated. The stage of the disease can be determined by examining CSF and observing trypomastigotes.

Figure 2.35 Photomicrograph showing two *Trypanosoma brucei* parasites in a blood sample. Courtesy of the Centers for Disease Control and Prevention.

Amoebiasis

Amoebiasis is an intestinal infection of *Entamoeba histolytica*. It is thought to cause 40–50 million new cases yearly and result in 40 000–100 000 deaths annually. Amoebiasis is acquired by ingesting faecally contaminated food or water. However, approximately 80% of infected people are asymptomatic and, of the unlucky remainder, most have only mild symptoms. These include diarrhoea, stomach pains and cramps. Infections are clinically most debilitating in the young and elderly. Amoebic dysentery is the severest form, associated with parasites invading the intestinal mucosa (*Chapter 13*). Its symptoms include abdominal pain, bloody stools, fever and vomiting. In rare cases, *E. histolytica* may cause invasive extraintestinal disease leading to peritonitis and liver abscesses. It can spread to the lungs and brain but, again, this is uncommon. Most fatalities are due to complications such as liver abscesses.

Figure 2.36 *Giardia lamblia* is named after the French biologist Alfred Giard (1846–1908).

Giardiasis

Giardiasis is an intestinal illness caused by the protozoan *Giardia lamblia* (synonyms *G. duodenalis* and *G. intestinalis*; *Figure 2.36*; see also *Figure 3.21*). The source of the infection is usually water contaminated with faecal material. Acute giardiasis develops

after 1–14 days and normally lasts one to three weeks. Infections can be asymptomatic but inflammatory disease presents with abdominal cramps, foul-smelling diarrhoea, fatigue, nausea and loss of weight. With chronic giardiasis, the symptoms reoccur and are debilitating, and food malabsorption may present.

Leishmaniasis

Leishmaniasis is an infection with protozoan *Leishmania* spp. (*Figure 2.37*). The vectors for these parasites are blood-sucking *Phlebotomus* sand flies (see *Figure 2.45*). Like other protozoan parasites, it undergoes several stages in development in its primary and secondary hosts. The three major types of leishmaniases are:

1. cutaneous leishmaniasis or oriental sore, occurring largely in the Middle East
2. visceral leishmaniasis or kala-azar, encountered most frequently in tropical Asian regions
3. mucosal or American leishmaniasis, found mainly in Central and South America

Some infected persons remain asymptomatic; however, the condition can develop in different ways dependent upon the type of leishmaniasis contracted. Cutaneous leishmaniasis is the most common type. It is so named from the sores that develop on the skin, usually several few weeks or even months after infection from a sand fly bite. The sores can change in size and appearance. They may begin as papules or nodules but can develop into skin ulcers that can scab or be covered with a crust. Normally they are painless but can be painful. Some patients have swollen glands near the sores. Visceral leishmaniasis is the other main form. It typically develops within months of the sand fly bite but can take years. Visceral leishmaniasis usually presents with fever, weight loss and enlargement of the liver and spleen. It also affects the bone marrow, resulting in anaemia, leukopenia and thrombocytopenia (*Chapter 13*). The condition can be life threatening. Mucosal leishmaniasis is caused by some *Leishmania* spp. that can spread from the skin to mucous membranes, where they can cause sores, commonly in the nose but also the mouth and throat. Hence, it may occur as a sequela of cutaneous leishmaniasis.

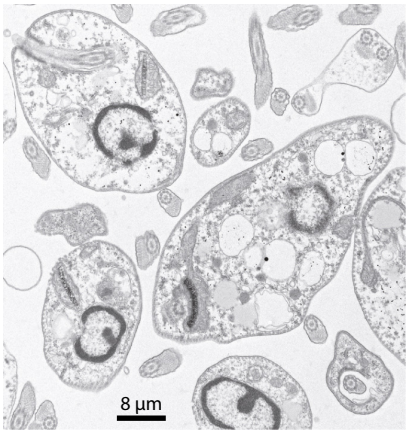


Figure 2.37 Electron micrograph of *Leishmania major* grown in a cell culture. The densely staining regions in the cytoplasm are kinetoplasts. Courtesy of the Centers for Disease Control and Prevention/Cynthia Goldsmith and Luciana Flannery. ⏎

Toxoplasmosis

Toxoplasmosis is caused by the parasite *Toxoplasma gondii*. The WHO has estimated that over one million cases occur annually in Europe. Indeed, toxoplasmosis is thought to be a leading cause of food-borne illness deaths in the USA. *T. gondii* is an intracellular parasite of cats, which become infected when they eat birds and mice with the parasite. It is transmitted to humans through ingesting soil, water or food, particularly undercooked meat, contaminated with cat faeces. People infected with *T. gondii* but with healthy immune systems (*Chapter 4*) are usually symptomless. When illness does occur, it is normally mild and presents with a range of symptoms, such as tender lymph nodes and muscle aches, and can be mistaken for influenza. These last for several weeks to months before subsiding; however, the person remains a carrier of the parasite, which can become reactivated if the person becomes immunosuppressed (*Figure 2.38*). Such patients can experience severe symptoms following necrosis in a range of tissues. These include CNS disorders, fevers, listlessness and a generalised malaise, lymphadenopathy, and inflammation of cardiac muscle and lung alveoli (myocarditis and pneumonitis, respectively).

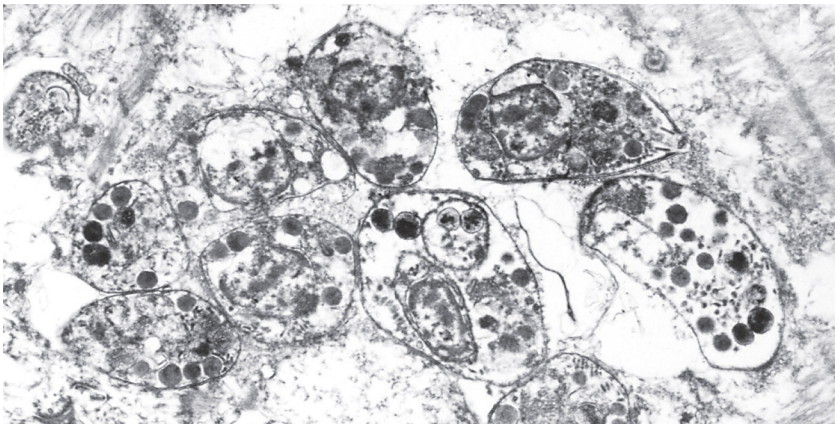


Figure 2.38 Electron micrograph of *Toxoplasma gondii* cysts in a biopsy sample of cardiac tissue from an immunocompromised patient with acquired immune deficiency syndrome (AIDS). Cysts can be as large as 60 µm in diameter. Courtesy of the Centers for Disease Control and Prevention/Dr Edwin P. Ewing Jr. ⏎

Toxoplasmosis can also occur by transplacental (congenital) transmission from a pregnant woman in the acute infected stage to the baby. Clinical presentation following congenital infection can be a miscarriage or stillbirth. Babies born with toxoplasmosis can suffer severe problems, such as cerebral damage with an abnormal accumulation of CSF (hydrocephalus) or small heads (microcephaly), intellectual disabilities, convulsions and blindness most frequently from retinochoroiditis.

Trichomoniasis

Trichomoniasis is caused by the protozoan *Trichomonas vaginalis* (see *Figure 3.32*). The WHO has estimated the global number of trichomonas infections at over 220 million patients. It also suggested its incidence in Europe is similar to that of the USA, where there were 2.6 million cases in 2018. In Africa, the prevalence of trichomoniasis is thought to be much higher. Trichomoniasis is a sexually transmitted infection and usually spreads during unprotected sexual intercourse. Its incubation time is 5–28 days, although up to 50% of infected people are asymptomatic carriers. In men and women, the initial symptoms include dysuria, pruritus (itching) and urethritis. In women, *T. vaginalis* mainly infects the vagina and urethra. This can result in an abnormal yellow-green-coloured discharge from the vagina that is often frothy in appearance and has an unpleasant fishy smell. Soreness, swelling and itching around the vulva may also present. In men, the infection commonly affects the urethra but also the head or foreskin of the penis. A thin, whitish-coloured discharge from the penis may also occur. Pain on micturition or during ejaculation can also be a symptom. The prostate gland can become infected. Complications of trichomoniasis are rare; however, without clinical treatment, although the discharges may cease, the condition can become chronic. Pregnant women infected with *T. vaginalis* may give birth prematurely to underweight babies.

2.9 FUNGI

Fungi are a heterogeneous group of organisms, ranging from unicellular yeasts to filaments of elongated chains of cells called hyphae, which are often called moulds (*Figure 2.39*; see also *Figure 3.15*).

Fungal infections, or mycoses, are relatively common and may be superficial or systemic. Mycoses are caused when fungi invade tissues, by their toxicity or by initiating allergic responses. Clinically important fungal infections include the dermatophytoses caused by the filamentous fungi *Epidermophyton*, *Microsporum* and *Trichophyton* spp. Ringworm exhibits the classic pattern of a dermatophytoses, being characterised by an inflamed circular lesion that is scaled with vesicles. These conditions are most common in hot, humid rural environments. Any general area of the skin may be infected (tinea corporis), including feet (tinea pedis or athlete's foot), groin (tinea cruris), scalp (tinea capitis) and toenails (tinea unguium; *Figure 2.40*).

Systemic fungal infections are, in general, much more common in immunocompromised individuals, for example those with AIDS (*Chapter 3*) or undergoing cancer chemotherapy (*Chapter 17*) and patients treated with immunosuppressive drugs to prevent rejection of a transplant (*Chapter 6*). The prognosis for systemic mycoses is often poor. Examples of such conditions include:

- aspergillosis
- candidiasis
- cryptococcosis
- histoplasmosis

Aspergillosis

Aspergillosis is an infection caused by fungi of the genus *Aspergillus*. Infections are relatively uncommon. Allergic bronchopulmonary aspergillosis (ABPA) has symptoms resembling asthma (*Chapter 5*). Chronic pulmonary aspergillosis leads to coughing, often with loss of blood, wheezing, fatigue and reduced weight. Invasive aspergillosis has similar symptoms but results in inflammation (*Chapter 4*) and granulomatous lesions on any organs.

Candidiasis

Candidiasis are infections, usually caused by *Candida albicans* (see *Figure 3.15*). Common manifestations include nappy (diaper) rash, thrush and vaginitis, where it is characterised by peeling of the covering membranes, bleeding and pruritus. Systemic forms may cause endocarditis and infections of the kidneys, liver, lungs and spleen,

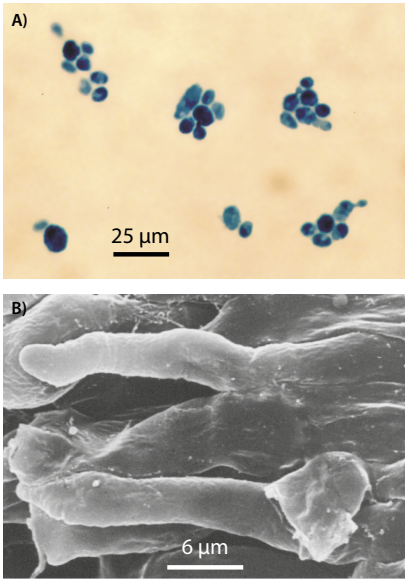


Figure 2.39 (A) Photomicrograph of *Saccharomyces cerevisiae* cells. (B) Electron micrograph of fungal hyphae. (A), Courtesy of M. Hoult, Nature Photography, www.naturalhistoryimages.co.uk. ⏎

Figure 2.40 Infections of the toenails (onychomycosis) with (A) *Trichophyton tonsurans* and (B) *Trichophyton rubrum* and secondary infections of the fungus *Scopulariopsis brevicaulis* and the yeast *Candida guilliermondii*. Courtesy of Dr Pavel Dubin, Israeli Board Certified Dermatologist. ⏎

which mainly occur in debilitated patients. Multidrug-resistant *Candida auris* is an increasing cause of candidiasis.

Cryptococcosis

Cryptococcosis is typically cause by inhalation of *Cryptococcus* spp., particularly *C. neoformans* and *C. gatti* (*Figure 2.41*). Infections are rare in healthy individuals but cause 100 000s of deaths annually, many in patients with advanced AIDS. These yeast-type organisms are found in dust, especially when contaminated with pigeon faeces. The infection can spread from the lungs to the CNS, skin and urinary system. Initial symptoms include coughing and associated respiratory problems. Later, nodules or tumours filled with gelatinous material develop in subcutaneous and visceral tissues. If the fungus infects the meninges, blurred vision, difficulties in speaking and headaches occur.

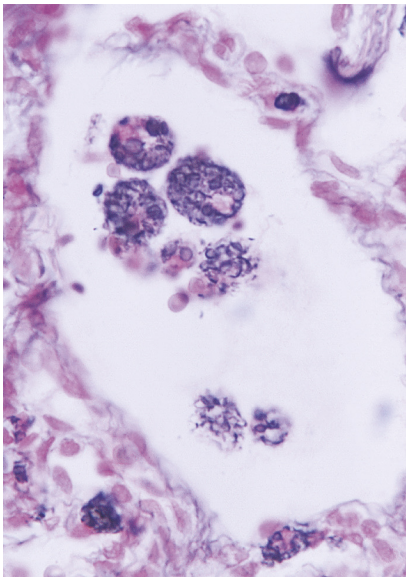


Figure 2.41 Photomicrograph showing pathological lung tissue associated with a *Cryptococcus* infection. Courtesy of the Centers for Disease Control and Prevention.

Histoplasmosis

Histoplasmosis or Darling disease was first described in 1906 by Samuel Darling (1872–1925). It is a major mycosis with an estimated 50 million people having latent infections and 500 000 new infections occurring annually. Histoplasmosis occurs extensively in the Americas and, indeed, is the most common endemic mycosis in the USA. It is caused by the dimorphic fungus *Histoplasma capsulatum*, which has a hyphal morphology when growing in soil. Disturbance of soil can release microconidia or hyphal fragments that, if inhaled, can settle in the alveoli of the lungs. Here, the fungus transforms into a single-cell form, reminiscent of the yeast *Saccharomyces cerevisiae* (see *Figure 2.39*). This change in morphology is thought to occur inside phagocytes, which can spread the pathogen initially to lymph nodes and later to multiple organs to give a disseminated histoplasmosis. Histoplasmosis affects both immunologically intact and deficient hosts, although the severest form is more likely in the latter. Approximately 90% of infections are asymptomatic, while localised, self-limiting lung infections present with influenza-like symptoms. About 1 in 100 000 cases progresses to a chronic pulmonary form with fibrosis and cavities forming at alveoli apices; of these patients, 1 in 2000 can progress to disseminated histoplasmosis with fevers, weight loss and a malaise that can be fatal if left untreated.

2.10 HELMINTHS

Helminths or parasitic worms are of clinical importance. Many have complex life cycles often involving hosts additional to humans, where they infect the GIT, blood vessels, lymphatics or other tissues such as skeletal muscle. Helminths are major parasites in tropical climates but some cause significant clinical problems in temperate climes. Poor sanitary conditions, contaminated food and water, and undercooked food are all associated with their transmission.

Helminths may be divided into two colloquially named groups:

1. flatworms (platyhelminths), comprising cestodes and trematodes
2. roundworms or nematodes

Flatworms

Cestodes include a variety of tapeworms, such as the pork tapeworm *Taenia solium* (see *Figure 3.64*), and are widespread common parasites. Infection occurs on eating food contaminated with faeces containing worm eggs or by eating raw or undercooked meat (pork in the case of *T. solium*) that contains the tapeworm or its larvae. Within the GIT, the larvae mature into long-lived hermaphroditic adult forms that can reach 3–5 metres in length. Release of eggs in faecal material allows them to be ingested by pigs and completion of the life cycle. Symptoms of intestinal infection are often mild but can include abdominal pains, diarrhoea, fatigue and weight loss. Infrequently, clinically serious encystations of muscles and the brain occur.

Parasitic trematodes include various liver flukes. Schistosomiasis, also called bilharzia or snail fever, is caused by *Schistosoma* spp. The WHO estimates that schistosomiasis affects approximately 240 million people globally, with several million suffering severe morbidity, and threatens more than 700 million people living in endemic areas. Indeed, only malaria affects more patients in the Middle East and tropical countries. The infection is prevalent in poor communities without potable water and adequate sanitation. It is contracted when the infective cercarial stage of schistosomes develop and emerge from freshwater snails, penetrate the skin and are transported in the blood. They

eventually mature into flukes that may live up to 20 years, living mainly in veins that drain the urinary tract and intestines, but they can affect other organs. On numerous occasions, they release eggs that become trapped in the tissues, causing obstruction, pain and organ dysfunction. Eggs released into freshwater can infect snails and promote the life cycle. Normally schistosomiasis is not fatal immediately, but it is a chronic illness that may seriously damage internal organs. In children, it is associated with reduced growth and cognitive development.

Roundworms

Ascariasis, an infection of tropical and temperate areas, is caused by the nematode *Ascaris lumbricoides* (*Figure 2.42*). It begins when *Ascaris* eggs are ingested and hatch in the small intestine. The resulting larvae pass through the intestinal wall and migrate in blood and lymph to the lungs. After 7–14 days, they enter the throat and are swallowed. The larvae mature in the jejunum to adult worms up to 300 mm long. These lay eggs that are passed out in faeces, contaminating soil and allowing completion of the life cycle. Early symptoms are coughing, wheezing and fever, while later ones include abdominal cramps and obstruction with diarrhoea and vomiting, fatigue and weight loss. In children, the migration of adult worms into the liver, gallbladder or peritoneum cavity may cause death.

Hookworm nematodes are parasites common in the tropics. They affect an estimated 740 million people worldwide. The major nematode culprit is *Necator americanus*, whose larvae live in soil and enter the patient through the skin. The larvae travel via the bloodstream to the lungs; from there they migrate to the oral cavity and are swallowed. In the jejunum, hookworms attach their mouths to the mucosal lining and feed on blood. Symptoms are abdominal pain and an iron-deficient anaemia.

Toxocara canis is a common roundworm found in dogs. It causes toxocariasis in humans, which is common in tropical countries and in Europe and North America. Infection starts with the ingestion of eggs in contaminated soil, particularly by children. Usually the eggs are eliminated by defaecation, but, in rare cases, they hatch, releasing larvae that spread throughout the body. Symptoms include an enlarged liver, eosinophilia, breathing problems, skin rashes and eye lesions that can lead to blindness.

Trichinosis is caused by the nematode *Trichinella spiralis*. It is transmitted in raw or undercooked meat infected with worm cysts. The action of stomach acid releases the larvae, which mature in the intestines into adult roundworms up to 3 mm long. Reproduction results in larvae that cross the intestinal mucosa and migrate in blood and lymph to skeletal muscle tissue where they form cysts (*Figure 2.43*). Early symptoms include abdominal pain, fever, nausea and vomiting; later ones are eosinophilia, fatigue and joint and muscle pains. The condition may be fatal if encystation occurs in the heart or brain.

Filarial worms, such as *Wuchereria bancrofti*, are common nematodes in tropical and subtropical regions. More than 200 million people are infected with filariasis: infectious conditions that include elephantiasis, heartworm and river blindness. Infection occurs when worm larvae enter the host through mosquito and blackfly bites. These develop into adult female worms 65–100 mm long but only 0.25 mm in diameter, with males even shorter and narrower, which particularly infest lymph ducts and nodes. Adult females produce large numbers of microscopic, active embryos called microfilariae that enter the bloodstream and infect insect vectors when they feed on the patient. Within the insect, the microfilariae develop into larvae, which may then infect other people when the insect feeds. Symptoms of filariasis include an increased susceptibility to other infections and damage to the lymph system. Blockage of lymphatic vessels leads to oedema distal to the obstruction with pain and swelling. After many years of infection, the end stage lesion is elephantiasis with grotesquely swollen external genitalia and legs with blackened and coarsened skin.

The nematode *Enterococcus vermicularis* is called a threadworm or human pinworm because of its pointed tail. It is one of the most common helminth infections in the world, with most cases occurring in children and, indeed, is the most common intestinal parasite of young children in the developed world. Infection may occur by ingesting eggs from contaminated surfaces, clothes or bed linen. Following ingestion, the eggs hatch in the small intestine, releasing larvae that develop into adult females 8–13 mm and males 2–5 mm long, respectively, which live mainly in the colon for about two months. Nocturnally, pregnant female worms migrate outside the anus and deposit eggs on perianal folds in the area. Self-inoculation can happen when eggs are transferred into the mouth after scratching the perianal area. The infection is frequently asymptomatic but particularly at night can cause pruritus of the perianal area. Secondary bacterial infections can result

Figure 2.42 *Ascaris lumbricoides* (a nematode). Courtesy of Shutterstock.com/Rattya Thongdumhya. ⏎

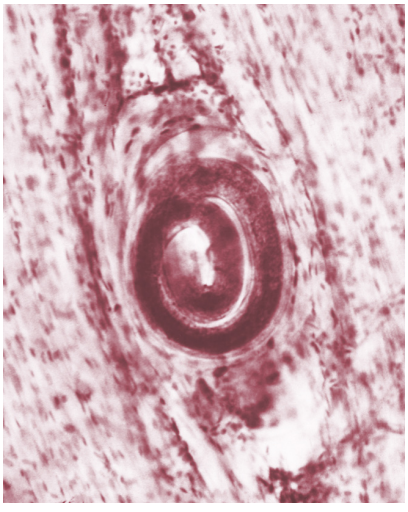


Figure 2.43 Photomicrograph of a *Trichinella spiralis* cyst embedded in a muscle tissue. Courtesy of the Centers for Disease Control and Prevention/Dr I. Kagan. ⏎

if excessive scratching damages the epithelial surfaces. The term threadworm is sometimes restricted to the pinworm *Strongyloides stercoralis*, which causes strongyloidiasis. This infection is acquired when soil-borne larvae penetrate the skin, causing a pruritic rash. The larvae are transported to the lungs in the bloodstream where they can cause pneumonia. Migration to the oral areas allows them to be swallowed and mature into adults in the small intestine. Symptoms include intestinal malabsorption with its attendant problems (*Chapter 11*) and bloody diarrhoea. A fatal disseminated strongyloidiasis occurs in rare cases.

2.11 ARTHROPODS AND VERTEBRATES

Houseflies and cockroaches (see *Figure 2.1*) are noted mechanical carriers of food-poisoning organisms. However, some arthropods are directly parasitic but many are also relevant to infectious diseases as vectors of pathogens. The mite *Sarcoptes scabiei* lives in the outer layers of the skin and can cause scabies. Blood-sucking parasites include human fleas and body lice (see *Figure 2.16*) and the pubic louse (*Pthirus pubis*). More specific vectors include ticks (*Ixodes* spp.) that transmit *Borrelia burgdorferi* (*Figure 2.44*), the cause of Lyme disease (*Chapter 3*), and *Y. pestis*, the plague organism, which uses human fleas and body lice as vectors (see *Box 2.2*). Malarial parasites are spread by female *Anopheles* mosquitos (see *Box 2.4*). Tsetse flies (*Glossina* spp.) are vectors of *T. brucei*, which causes sleeping sickness, while *Phlebotomus* sandflies are vectors of *Leishmania* spp., the causative agents of leishmaniasis (*Figure 2.45*). Infection with *T. cruzi* is spread by kissing bugs and causes Chagas disease (*Box 2.5*). Several disease-causing organisms use vertebrates, such as bats or mammals, as vectors with perhaps the best known being the rabies virus.

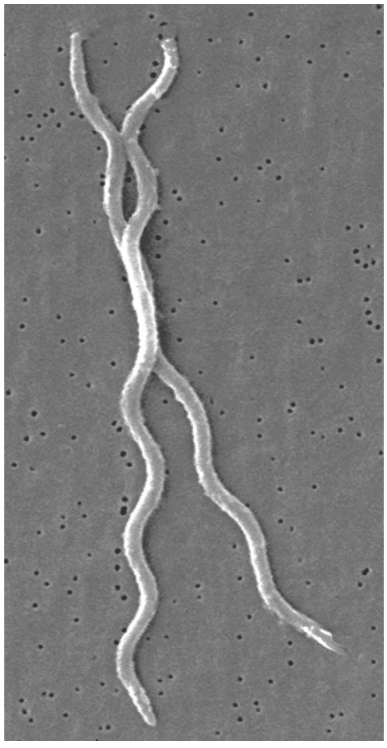


Figure 2.44 Electron micrograph of *Borrelia burgdorferi*. Courtesy of the Centers for Disease Control and Prevention/J.H. Carr.

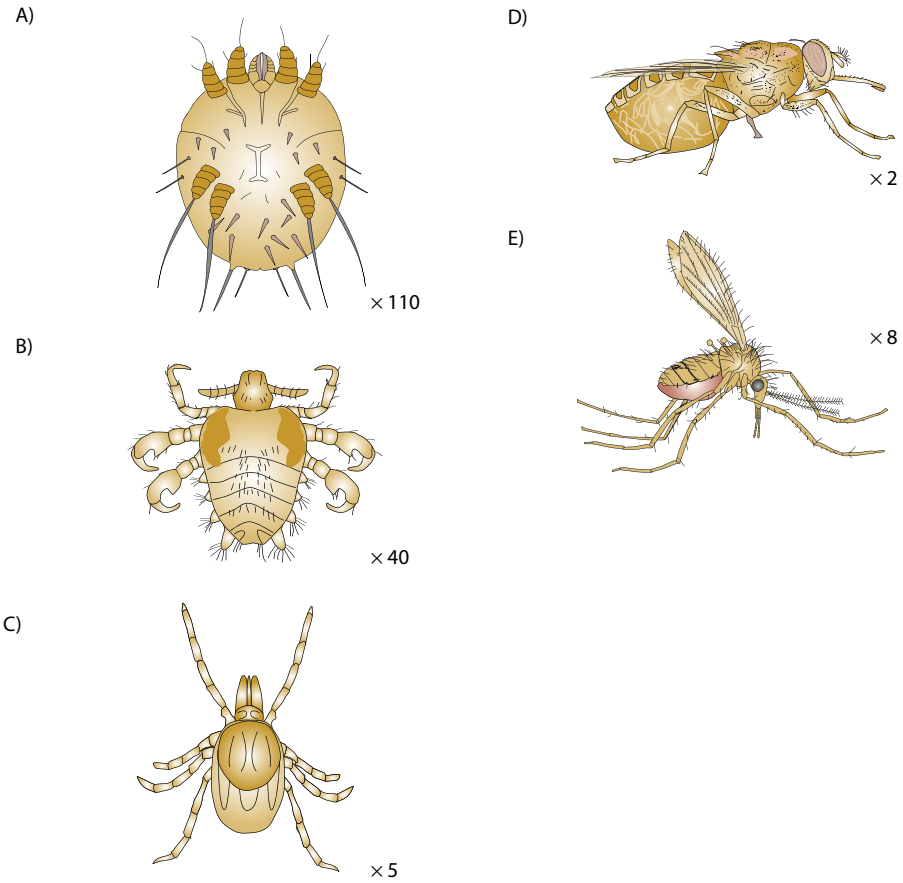


Figure 2.45 Examples of arthropod parasites and vectors. (A) Scabies mite (*Sarcoptes scabiei*), (B) pubic louse (*Phthirus pubis*), (C) tick (*Ixodes ricinus*), (D) tsetse fly (*Glossina* sp.) and (E) sand fly (*Phlebotomus* sp.).

BOX 2.5 A NEGLECTED TROPICAL DISEASE – CHAGAS DISEASE

Neglected tropical diseases (NTDs) are so-called because they receive little attention in terms of the global health agenda and limited resources and funding from global agencies, resulting in the diverse group listed by the World Health Organization (WHO; *Table 2.6*). More than one billion people worldwide suffer from NTDs. They mainly live in impoverished conditions and, by definition, in tropical areas with poor sanitation and access to uncontaminated water. Thus, NTDs are diseases of neglected populations and contribute to cycles of poor educational outcomes and limited career opportunities. Neglected tropical diseases are caused by a variety of pathogens, including viruses, bacteria, protozoa and fungi. Many of the pathogens associated with NTDs are vector-borne with complex life cycles and have extensive animal reservoirs, factors that make their control challenging. Chagas disease exemplifies many of these points.

Buruli ulcer
Chagas disease (American trypanosomiasis)
Dengue and chikungunya
Dracunculiasis (Guinea-worm disease)
Echinococcosis
Food-borne trematode infections
Leprosy (Hansen disease)
Leishmaniasis
Lymphatic filariasis (elephantiasis)
Onchocerciasis (river blindness)
Schistosomiasis (bilharzia)
Rabies
Scabies and other ectoparasitoses
Snakebite envenoming
Taeniasis and cysticercosis
Soil-transmitted helminthiases
Trachoma
Yaws (endemic treponematoses)

Table 2.6 Neglected tropical diseases (NTDs) listed by the World Health Organization.

Chagas disease is named after the Brazilian physician Carlos Chagas (1879–1934), who discovered the disease in 1909. In 2019, it was estimated that over six million people living mainly in rural, poor regions of Mexico and Central and South America had the disease, where it causes approximately 9 500 deaths annually. Unfortunately, large-scale population migrations and widespread intercontinental travel have carried it to the USA and many European countries.

The disease is caused by *Trypanosoma cruzi*, hence its alternative name of American trypanosomiasis. The vectors of this are parasite blood-sucking *Triatoma* insects, colloquially called 'kissing bugs' (*Figure 2.46*), which spread the infection when feeding. An infected insect releases the parasites in its faeces near the site of the bite wound when taking a blood meal, which allows them to enter the host. Like other protozoal parasites, such as *Plasmodium* spp., *T. cruzi* has a complicated life cycle, passing through several stages in both humans and kissing bugs.

Figure 2.46 A *Triatoma* kissing bug. Courtesy of Dr D. Dörge, Goethe University, Frankfurt, Germany.

Chagas disease presents with acute and chronic phases. The former is normally asymptomatic but can be associated with general somatic symptoms. In rare cases, it may present severe potential cardiac or neurological symptoms and signs. Nodular lesions, usually called chagomas, can develop at the feeding site of the bug. Most acute cases resolve over a period of weeks or months into an indeterminate or subclinical chronic form of the disease. The dormant asymptomatic phase can be reactivated if the patient becomes immunocompromised. However, the patient may remain in the indeterminate state for years or decades after the initial infection. Indeed, most people do not realise they are infected before developing symptoms of the chronic form. Many of the complications of chronic Chagas disease, which may include simultaneous cardiac or GIT problems, can be fatal.

In the acute phase, Chagas disease can be diagnosed by the presence of *T. cruzi* in blood samples (*Figure 2.47*). It is more difficult in the chronic stage, when the organism has migrated from blood to other organs. It may be observed in biopsy material from bone marrow, liver, lymph nodes or spleen. Treatment is prophylactic, including sealing houses against kissing bugs and educating potential victims of their danger. Drugs to treat the acute phase include allopurinol, benznidazole and nifurtimox. Treatment at the chronic stage is more problematic, since eradicating the parasite is more difficult and, indeed, may be impossible. Again, treating with benznidazole or nifurtimox for chronic Chagas disease patients under the age of 55 years who do not have advanced irreversible cardiomyopathy is recommended.

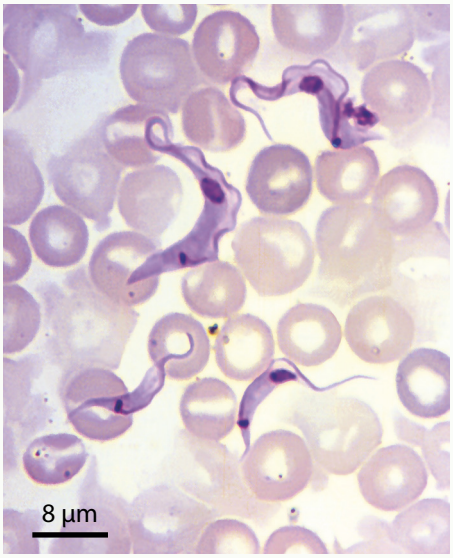


Figure 2.47 Photomicrograph showing four flagellated *Trypanosoma cruzi* in a blood sample. Courtesy of the Centers for Disease Control and Prevention.

2.12 CASE STUDIES

CASE STUDY 2.1

A 38-year-old male, Brian, was admitted into hospital two weeks after spending a year working in several African countries. Seven days before admission, Brian had developed coughing, muscular pains and recurrent chills and fevers approximately three times daily. He was also slightly jaundiced.

Questions

(a) What is the most likely disease affecting Brian?
(b) Which organism(s) cause(s) this disease? How could this be confirmed?
(c) How could this disease have been prevented?

CASE STUDY 2.2

Anthony, a 62-year-old engineer, had to make a trip at very short notice to a developing country to oversee the repairs of essential machinery. Despite being unvaccinated, he agreed and visited several rural populations. The day after returning, he visited his general practitioner (GP) complaining of a sore throat and headaches. He was advised to rest and drink 'plenty of fluids'. Two days later, he attended the accident and emergency department of his local hospital with nausea and some difficulties in swallowing and a deterioration in his sore throat. Anthony was prescribed a short dose of antibiotic tablets and advised to see his GP. Three days later, he was admitted into hospital with breathing difficulties. Examinations showed an enlarged epiglottis, yellow-coloured exudates on his oral structures and fluid penetrating his lungs. Despite antibiotic therapy, his condition deteriorated and he was transferred to the intensive care unit. Within a week, fever had developed and chest radiography showed the bases of his lungs to be heavily infected: symptoms typical of diphtheria. This was confirmed when clinical specimens grew colonies of *Corynebacterium diphtheriae* on tellurite blood agar (*Figure 2.48*). Unfortunately, despite an aggressive drug regimen and the use of antidiphtheria toxin, Anthony's condition continued to worsen, and he eventually died of heart and kidney failures.

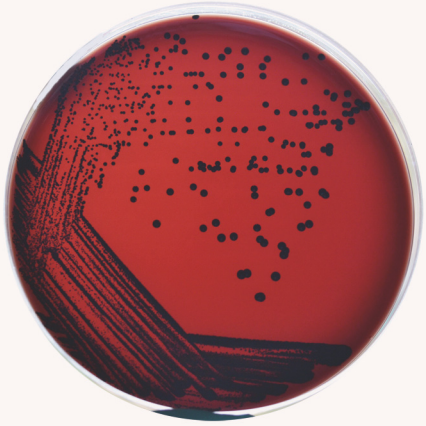


Figure 2.48 Black-coloured colonies of *Corynebacterium diphtheriae* growing on tellurite blood agar. Courtesy of Shutterstock.com/Chatchouliya.

Question

(a) What do you consider to be the important aspects of this case study?

CASE STUDY 2.3

A 77-year-old man, Douglas, returned to the UK after living in South India for three years. He was admitted to hospital complaining of dizziness, shortness of breath, some abdominal pain and general weakness over the last two years. On examination, he was pale and appeared ill and severely undernourished but otherwise his vital signs were normal. His blood sample gave the following results (reference ranges in parentheses):

Erythrocyte count:	1.8×10^{12}/L	($4.4–5.9\times10^{12}$/L)
Haemoglobin:	35 g/L	(140–180 g/L)

A blood smear showed his erythrocytes were microcytic and hypochromic. Stools were negative for occult blood. Endoscopy of the duodenum showed the presence of reddish-coloured slender moving roundworms approximately 10 mm in length that were subsequently identified as the hookworm *Necator americanus*.

Questions

(a) Discuss the clinical, blood and stool results.
(b) How should Douglas be treated?

2.13 SUMMARY

Diseases in humans may be caused by prions and viruses, and organisms such as bacteria, protozoa, fungi and helminths. Prions are aberrant proteins that can be transmitted in, for example, food. Viruses are macromolecular complexes; all are obligate intracellular parasites. A number of microorganisms are not normally pathogenic but can become so when a new 'niche' becomes available to them, for example a burn wound, or when the host is immunocompromised in some way. Bacterial pathogens display a variety of virulence factors that aid entry into the host and help them to overcome host defences. These virulence factors can have a generalised action or be more specifically defensive or offensive in nature. Examples of offensive virulence factors include the production of endo- and exotoxins. Some protozoan pathogens possess a hypermutable genome that enables them to change the nature of their surface antigens and so escape the host's immune system. Fungi and helminths are eukaryotic parasites responsible for many clinical conditions and deaths worldwide. A number of arthropods and some vertebrates are parasitic, and some are significant vectors in the spread of several infectious diseases.

QUESTIONS

1. To which class of parasite or pathogen do the following belong?
Trypanosoma brucei, HIV, *Microsporum* species, *Treponema pallidum*, *Candida albicans*.

2. Arrange the following two lists into their most appropriate pairings.

Clostridium perfringens	Shigellosis
Epstein–Barr virus	Gas gangrene
HIV	Chicken pox
Human rhinovirus 91	Bubonic plague
Microsporum and *Trichophyton* species	CD4
Plasmodium	Malaria
Sarcoptes scabiei	Glandular fever
Shigella dysenteriae	Scabies
Varicella zoster	Intracellular adhesion molecule
Yersinia pestis	Athlete's foot

3. Compare and contrast the actions of endo- and exotoxins.

4. Which of the following is the odd ONE out?
(a) Human flea (*Pulex irritans*).
(b) Human body louse (*Pediculus humanus*).
(c) Tsetse fly (*Glossina* spp.).
(d) Housefly (*Musca domestica*).
(e) Kissing bug (*Triatoma* spp.).

FURTHER READING

Belete TM (2020) Recent progress in the development of new antimalarial drugs with novel targets. *Drug Des Dev Ther* **14**:3875–3889.

Bern C (2015) Chagas' disease. *N Engl J Med* **373**:456–466.

Bhadoria P *et al*. (2021) Viral pandemics in the past two decades: an overview. *J Fam Med Prim Care* **10**:2745–2750.

Caughey B (2001) Interactions between prion protein isoforms: the kiss of death? *Trends Biochem Sci* **26**:235–242.

Dean KR *et al*. (2018) Human ectoparasites and the spread of plague in Europe during the second pandemic. *Proc Natl Acad Sci U S A*, **115**:1304–1309.

DeArmond SJ and Bouzamondo E (2002) Fundamentals of prion biology and diseases. *Toxicology* **181–182**:9–16.

Demeure CE *et al*. (2019) *Yersinia pestis* and plague: an updated view on evolution, virulence determinants, immune subversion, vaccination, and diagnostics. *Genes Immun* **20**:357–370.

Gill ON *et al*. (2020) Prevalence in Britain of abnormal prion protein in human appendices before and after exposure to the cattle BSE epizootic. *Acta Neuropathol* **139**:965–976.

Glatzel M *et al*. (2005) Human prion diseases: molecular and clinical aspects. *Arch Neurol* **62**:545–552.

Global Health Metrics (2020) Chagas disease – level 3 cause. *Lancet*. Archived PDF from the original on 18 March 2022 (available at https://web.archive.org/web/20220318033609/https:/www.thelancet.com/pb-assets/Lancet/gbd/summaries/diseases/chagas-disease.pdf).

Green LH and Goldman E (2021) *Practical Handbook of Microbiology*, 4th edn. Taylor & Francis.

Guarner J (2019) Chagas disease as example of a reemerging parasite. *Semin Diagn Pathol* **36**:164–169.

International Committee on Taxonomy of Viruses Executive Committee (2020) The new scope of virus taxonomy: petitioning the virosphere into 15 hierarchical rank. *Nat Microbiol* **5**:668–674.

Lidani KCF *et al*. (2019) Chagas disease: from discovery to a worldwide health problem. *Front Public Health* **7**: Article 166. doi: 10.3389/fpubh.2019.00166.

Loker ES and Hofkin BV (2022) *Parasitology: A Conceptual Approach*, 2nd edn. Taylor & Francis, Garland Science.

Madigan MT *et al*. (2021) *Brock Biology of Microorganisms*, 15th edn. Pearson Education.

Magner LN and Kim OJ (2017) *A History of Medicine*, 3rd edn. Taylor & Francis.

Mittal J *et al*. (2019) *Histoplasma capsulatum*: mechanisms for pathogenesis. *Curr Top Microbiol Immunol* **422**:157–191.

Mortlock S (2023) Death and disease amongst the rich and shameless. *Biomed Sci* (December): 26–31.

Murray PR *et al*. (2020) *Medical Microbiology*, 9th edn. Elsevier.

Pérez-Molina JA and Molina I (2018) Chagas disease. *Lancet* **391**:82–94.

Senthilingam M (2020) *Outbreaks and Epidemics*. Icon Books Ltd.

Sharp PM *et al*. (2020) Ape origins of human malaria. *Annu Rev Microbiol* **74**:39–63.

Shaw-Taylor L (2020) An introduction to the history of infectious diseases, epidemics and the early phases of the long-run decline in mortality. *Econ Hist Rev* **73**:E1–E19.

Struthers JK (2017) *Clinical Microbiology*, 2nd edn. CRC Press.

Tao Y-X and Conn PM (2018) Pharmacoperones as novel therapeutics for diverse protein conformational diseases. *Physiol Rev* **98**:697–725.

Tse EG *et al*. (2019) The past, present and future of anti-malarial medicines. *Malar J* **18**:93. doi: 10-1186/s12936-0192724-z.

van Seventer JM and Hochberg NS (2017) *Chapter 6*. Principles of infectious diseases: transmission, diagnosis, prevention, and control. In: Quah SR (ed.), *International Encyclopedia of Public Health*, Elsevier, pp. 22–39.

Wicht KJ *et al*. (2020) Molecular mechanisms of drug resistance in *Plasmodium falciparum* malaria. *Annu Rev Microbiol* **74**:431–454.

Useful websites

www.ecdc.europa.eu
www.cdc.gov
www.cjd.ed.ac.uk
www.who.int
https://www.who.int/news-room/questions-and-answers/item/neglected-tropical-diseases

Prevention, control and treatment of diseases

3

OBJECTIVES

After studying this chapter you should be able to:

- describe the types and symptoms of some of the most common infectious diseases
- outline measures used to prevent infectious diseases
- outline measures used to control infectious diseases
- discuss general ways in which infectious diseases are investigated to aid diagnoses
- list the general effects of antibiotics on infectious organisms
- review the methods in which infectious organisms adapt to resist therapeutic drugs
- discuss some general aspects of the management and treatment of specific infectious diseases

3.1 INTRODUCTION

Pathogens and parasites (*Chapter 2*) cause a range of infectious diseases (*Figure 3.1*). A pathogen or parasite must be transmitted from a source to the patient. The sources of pathogens can be abiotic (soil or water for example) or animals or other humans (*Chapter 2*). Direct contact between hosts is the most obvious form of transmission, but coughs and sneezes (aerosols), food, water and arthropod vectors are all used by various pathogens. The long-term survival of pathogenic organisms also depends on them maintaining their infectivity during transmission from host to host.

In a text of this length, it is simply not possible to describe all infectious diseases. This chapter will therefore concentrate on selected, representative examples of infectious conditions and their causative organisms. In addition, the ways in which infectious diseases are prevented, controlled, investigated and treated will be outlined.

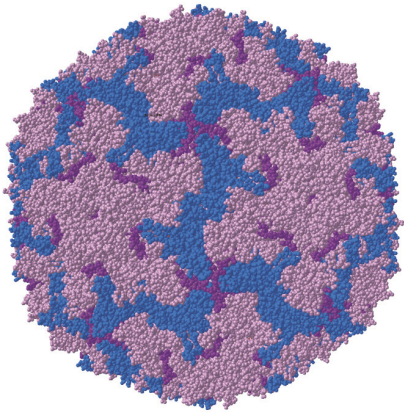


Figure 3.1 Molecular model of a parasitic Coxsackie virus, which is responsible for increasing numbers of severe infections in children. PDB file 6SMG.

3.2 INFECTIONS OF THE SKIN

The skin is a major element of innate immune defence (*Chapter 4*). It is normally colonised by a variety of microorganisms, although the numbers and types vary between different areas of the body. In normal circumstances, it forms an effective barrier to invading microorganisms.

Papillomaviruses can infect epidermal cells and stimulate their proliferation to form warts. Numerous herpes viruses have been described and, largely because of historical reasons, they are also known by other names (*Table 3.1*). At least eight of them, the **human herpes viruses** (HHV-1–8; *Figure 3.2*), infect humans and can cause a variety of clinical conditions. For example, HHV-1 causes herpes labialis. This is a common and often recurrent condition. The virus is spread directly by kissing or indirectly by sharing drinking utensils or toothbrushes. Herpes labialis is associated with lesions around and in the mouth, often called 'cold sores' or 'fever blisters'. The lesions begin as clear vesicles with abnormally red bases that develop into benign but painful ulcers that acquire a 'crusted' surface if untreated.

DOI: 10.1201/9781003296836-3

Human herpes viruses	Historical name
HHV-1	Herpes simplex virus 1 (HSV-1)
HHV-2	Herpes simplex virus 2 (HSV-2)
HHV-3	*Varicella zoster* virus (VZV)
HHV-4	Epstein–Barr virus (EBV)
HHV-5	Cytomegalovirus (CMV)
HHV-6	B lymphotropic virus (B-LV)
HHV-7	T lymphotropic virus (T-LV)
HHV-8	Kaposi sarcoma virus (KSV)

Table 3.1 Nomenclatures of the human herpes viruses (HHV-1–8). See also *Table 2.5*.

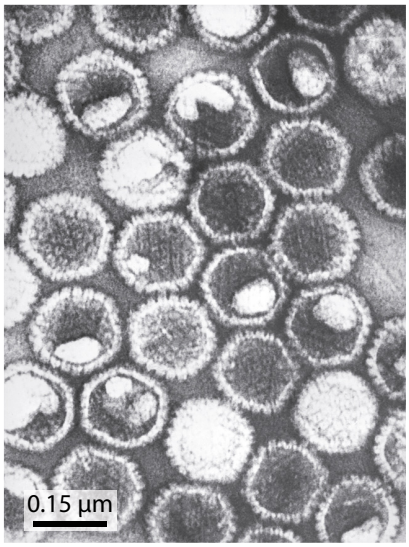


Figure 3.2 Electron micrograph of a human herpes virus (HHV). See also *Tables 2.5* and *3.1*. Courtesy of the Centers for Disease Control and Prevention/Dr F. Murphy and S. Whitfield.

Several antiviral drugs, for example acyclovir (*Box 3.1*), are available to treat the condition. Not all the herpes viruses are as relatively benign as HHV-1 and -2. *Varicella zoster* causes chicken pox and shingles (see *Figure 2.5*), and others are associated with several cancerous conditions (see *Table 3.1* and *Chapter 17*).

Bacterial skin infections normally occur only if the usual balance between the skin environment and these organisms is disturbed. The outbreaks of acne caused by *Propionibacterium* species (*Figure 3.5*) during the hormonal changes associated with puberty are typical. Breaks in the skin from wounds, surgery or lesions from insect bites or chicken pox may also lead to infections of the skin by *Staphylococcus aureus* (*Figure 3.6*) and *Streptococcus pyogenes*. In children especially, these bacteria may cause impetigo

BOX 3.1 ACYCLOVIR: THE BEST KNOWN ANTIVIRAL?

Acyclovir is used to treat human herpes virus (HHV)-1 and -3 infections. It is readily available as a proprietary preparation to apply topically to cold sores; for this reason, it is probably the best-known antiviral drug. Acyclovir is a structural analogue of deoxyguanosine (*Figure 3.3*). It is, perhaps, better described as a prodrug, since the viral thymidylate kinase (*Figure 3.4*) and host cell enzymes phosphorylate it, forming a substrate for the herpes DNA polymerase. During viral replication, the incorporation of the active drug into the growing DNA strand prevents its further elongation and leads to the cessation of viral growth.

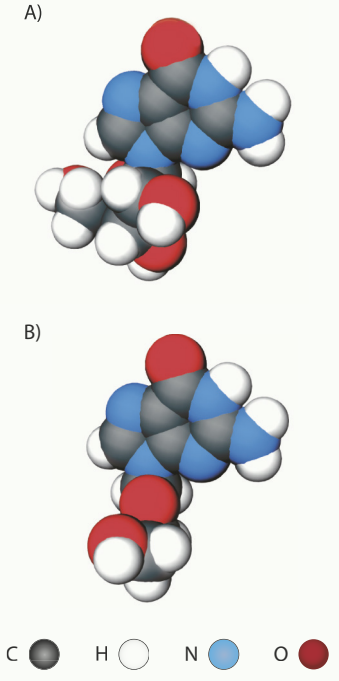


Figure 3.3 Computer drawn models of (A) deoxyguanosine and (B) acyclovir. Note the similarity in their structures.

Figure 3.4 Molecular model of thymidylate synthase from herpes virus 1, with the bound acyclovir molecules highlighted in red. PDB file 2KI5.

contagiosa, an extremely contagious skin infection. *S. aureus* can colonise hair follicles, leading to inflammations that can develop into abscesses (boils) or even, in extreme cases, **carbuncles**: an amalgam of several abscesses. Enterobacteria, *Pseudomonas aeruginosa*, *S. aureus* and *S. pyogenes* are all associated with a variety of skin problems following weakening of the patient, for example by diabetes (*Chapter 7*), deficiencies of the immune system (*Chapter 5*), **nosocomial** (hospital-acquired) infections of surgical wounds or burning of the skin. *Mycobacterium leprae* is the causative organism of the once-dreaded disease leprosy, which, although now virtually eliminated in developed countries, still affects many thousands in developing countries.

Fungi are associated with a variety of infections, particularly of the skin (see *Chapter 2* and *Figure 2.11*).

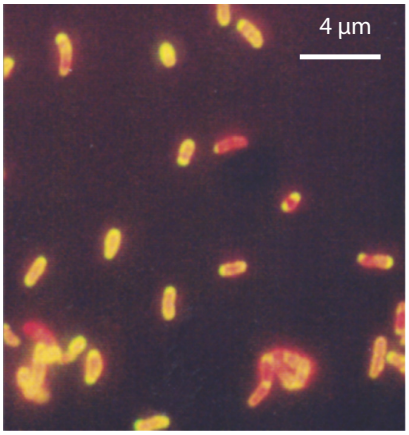


Figure 3.5 *Propionibacterium acnes*, a bacterium associated with skin infections such as acne. Courtesy of Dr S. Patrick, Queen's University Belfast, UK. ⏎

3.3 INFECTIONS OF SENSE ORGANS AND THE CENTRAL NERVOUS SYSTEM

All the major sense organs and the central nervous system (CNS) are subject to diseases arising from infectious pathogens and parasites.

Eyes

Eyelid infections generally involve the lid margins, eyelid glands or follicles, causing styes or **hordeolums**. They are usually associated with *S. aureus* infections. The conjunctiva is particularly susceptible to infection. The epithelial surface enclosed by the eyelids is a warm, moist environment in which microorganisms can rapidly become established. However, microorganisms must avoid being rinsed away by tears, and some pathogens, such as *Chlamydia trachomatis* (*Figure 3.7*), attach specifically to conjunctival cells. An estimated 500 million people are infected with different serotypes of *C. trachomatis*, making trachoma the most significant eye infection worldwide. The disease blinds approximately 1% of infected individuals, while many others suffer visual impairment. *C. trachomatis* is transmitted by contact with contaminated flies, fingers and towels, although trachoma itself results from chronic repeated infections. This is much more likely to occur in regions where restricted access to water prevents regular washing of the hands and face. Chlamydia is also a sexually transmitted disease (*Section 3.6*) and there is evidence that untreated chlamydial infections can lead to premature delivery, and babies born to affected mothers can have infected eyes and respiratory tracts. Chlamydia is a leading cause of conjunctivitis (pink eye) in newborns and early-infant pneumonia.

A laboratory diagnosis of trachoma can be carried out using samples of conjunctival fluid or scrapings. The usual treatment is with oral or topical antibiotics, such as tetracycline or doxycycline (*Section 3.11*). *Chlamydia* infections account for only 20% of cases of conjunctivitis; others are caused by bacteria such as *Streptococcus pneumoniae* and *Leptospira* spp.

Figure 3.6 Electron micrograph of *Staphylococcus aureus*. Courtesy of Dr A. Curry, Manchester Royal Infirmary, UK. ⏎

Serious infections of the inner eye with *P. aeruginosa* may follow trauma or after invasion by the protozoan *Toxoplasma gondii*, causing chorioretinitis and possible blindness. This widespread protozoan is not a serious threat unless acquired *in utero* or when an individual is immunosuppressed, perhaps as a result of taking drugs to prevent transplant rejection (*Chapter 6*). Infection occurs by swallowing oocysts released by infected cats or by eating meat containing cysts.

Eyes may also be infected by parasitic helminths ('worms'). For example, the tapeworm *Echinococcus granulosus* is transmitted when eggs are acquired from infected dogs. Infection by the larvae of the nematode helminth *Toxocara canis*, which occurs naturally in the intestine of dogs, is, however, more common. An infection can lead to the detachment of the retina. Onchocerciasis or 'river blindness' occurs in Africa and South America and is caused by the helminth *Onchocerca volvulus*. *Simulium* flies carrying larvae obtained from the skin of infected hosts transmit the infection. These flies develop in rivers, hence their colloquial name. River blindness is a serious infection, with over 18 million people affected worldwide and 270 000 blinded annually. The rates of blindness, which is irreversible, may reach 40%. The usual treatment is with anthelmintic drugs (*Section 3.11*).

Ears

Infections of the outer ear may cause pain or irritation and are resistant to some antibiotics. The middle ear can be colonised by viruses and bacteria from the upper

respiratory tract, causing acute otitis media with swelling and blockage of the Eustachian tube, which may lead to deafness, although this is generally temporary. Microorganisms that cause middle-ear infections include the mumps virus. This may be followed by secondary infections with *S. pneumoniae* and *Haemophilus influenzae*. Otitis media is the second most common paediatric condition after upper respiratory infections and particularly affects infants 6–24 months old. The general symptoms, apart from a devastating earache, are fever, vomiting and diarrhoea. The vesicles of the tympanum become dilated, with bulging of the drum itself occurring in the later stages of infection. If treatment is inadequate, then acute attacks may eventually perforate the eardrum, producing chronic discharges and defective hearing. The usual treatment is with oral antibiotics, such as ampicillin, amoxicillin, erythromycin and cefixime (*Section 3.11*).

Central nervous system

Infections of the CNS may affect the meninges, the spinal cord or the brain, causing meningitis, myelitis and encephalitis, respectively. More than one area may be infected simultaneously. Such infections can also become systemic (*Section 3.7*). Pathogens can enter these areas following head injuries, along the axons of neurons or by breaching the blood–brain or blood–cerebrospinal fluid barriers (*Figure 3.8*). The most common causes of viral meningitis are enteroviruses, such as enteric cytopathic human orphan (ECHO), Coxsackie (see *Figure 3.1*) and, formerly, poliomyelitis viruses. Viral meningitis is not usually a life-threatening condition. Bacterial meningitis, in contrast, has a mortality greater than 10%. The principal causative organisms are the capsulated bacteria *Neisseria meningitidis* (see *Figure 1.3*), *S. pneumoniae* and, before the introduction of a vaccine, *Haemophilus influenzae* (*Margin Note 3.1*).

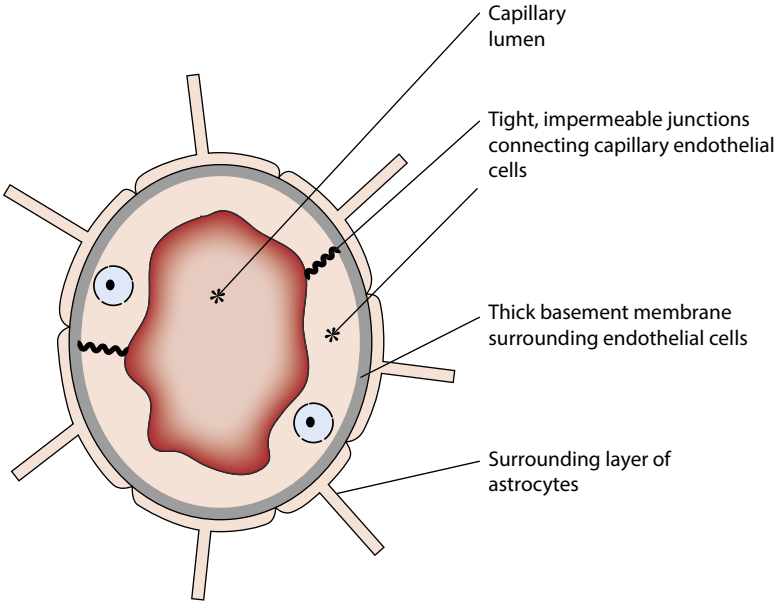

Capillary lumen
Tight, impermeable junctions connecting capillary endothelial cells
Thick basement membrane surrounding endothelial cells
Surrounding layer of astrocytes

Figure 3.8 Schematic of the blood–brain barrier. Note the tightly associated endothelial cells and thick basement membrane, which prevent most materials in the blood entering the brain.

Effective vaccines are also available against some serogroups of *N. meningitidis*, and a vaccine against *S. pneumoniae* is recommended by the Centers for Disease Control and Prevention (CDC) for use with children under five years old and all adults over the age of 65 years.

Encephalomyelitis results from infections by a number of viruses or protozoa. These include some poliovirus types, herpes simplex, measles virus, human immunodeficiency virus (HIV), toga viruses (which are transmitted by arthropods) and the rabies virus from infected mammals. The protozoan *T. gondii* may infect individuals with compromised immune systems, while *Trypanosoma brucei* is the causative organism of African sleeping sickness (*Chapter 2*).

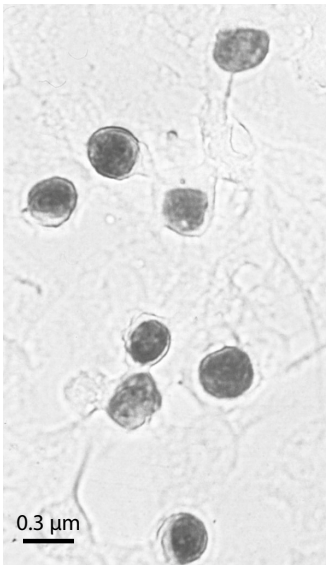


Figure 3.7 Photomicrograph of *Chlamydia trachomatis* growing in cultured eukaryotic cells. Courtesy of the School of Biochemistry and Microbiology, University of Leeds, UK.

Margin Note 3.1
***Haemophilus influenzae* type b**

Haemophilus influenzae type b is the cause of 'Hib' meningitis. It is also responsible for childhood epiglottitis, causing the throat to swell alarmingly, breathing difficulties as mucus collects in the throat, and fever. The condition is life threatening. *H. influenzae* type b can also cause pneumonia and other lower respiratory infections. The health risks are mainly associated with children under five years old, but adults whose resistance has been weakened by sickle cell anaemia (*Chapter 13*), chronic disease of the spleen, alcoholism (*Chapter 12*) or some malignancies (*Chapter 17*) are also at risk. The introduction of effective vaccines in the 1980s has significantly reduced Hib infections and their associated morbidity and mortality worldwide. The initial primary dose generally begins at six to eight weeks of age. Vaccine schedules vary between different countries, for example two primary doses followed by a single booster or three primary doses without a booster. However, all variations appear equally effective. The World Health Organization (WHO) recommends that Hib vaccines be included in routine infant immunisation programmes (see *Box 3.4*). It estimates that the Hib vaccination global coverage was 71% in 2021, but wide regional variations occurred. The future reduction of Hib infections via childhood vaccination programmes is a high priority.

3.4 INFECTIONS OF THE RESPIRATORY SYSTEM

The respiratory system is constantly exposed to inhaled microorganisms but is protected by extensive defences. The nose filters out particles larger than 10 μm, although those smaller than 5 μm may reach the bronchi and alveoli. In addition, there is a host of immune defences, including alveolar macrophages, secretory immunoglobulin A (IgA) antibodies, complement proteins, surfactant proteins, secreted defensins and lactoferrin (*Chapter 4*). Despite this, infections of the respiratory tract are frequent. The WHO has reported that many hundreds of millions of patients suffer acute infections of the lower respiratory tract worldwide. *Figure 3.9* indicates the sites of some respiratory diseases.

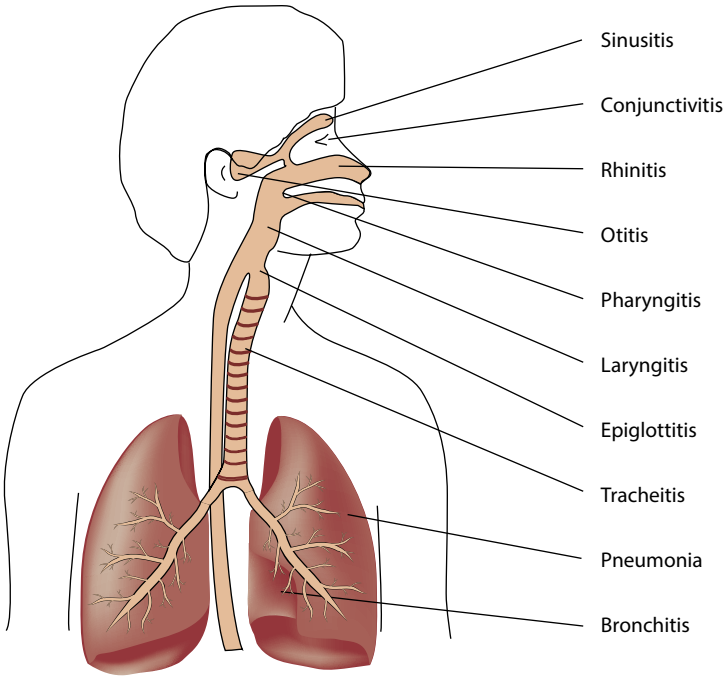


Figure 3.9 Examples of infections associated with the head and respiratory system. ⏎

Respiratory viruses are transmitted directly by aerosols or indirectly from contaminated surfaces. The first site of attack is, not surprisingly, epithelia of the nose and throat. Indeed, the hundreds of corona and rhinoviruses (*Figure 3.10*) that cause the common cold replicate at 32–33°C, the temperature of the mucosal surface lining the nose. The influenza viruses (see *Figures 2.10* and *2.14*) infect and replicate in respiratory epithelial cells causing cellular damage. The generalised symptoms that present, such as muscular aches, malaise and anorexia, suggest the virus may spread systemically from the respiratory tract, but there is no conclusive evidence for this.

The loss of ciliated and mucus-producing epithelial cells impairs clearance of invading microbes and creates conditions for secondary bacterial infections of staphylococci, streptococci or *H. influenzae.* Bacterial proteases, for example the V8 protease of *S. aureus*, can enhance the infectiveness of the influenza virus by improving virus adhesion.

Bordetella pertussis and *Corynebacterium diphtheriae* (*Figure 3.11*) are obligate bacterial pathogens. *C. diphtheriae* infects the nasopharynx and the tonsils and can progress to a lethal systemic infection affecting the heart, liver and kidneys. *H. influenzae* can infect the epiglottis, while *B. pertussis* is the causative agent of whooping cough. It adheres to the epithelial cells lining the trachea and bronchi, where it interferes with ciliary action and releases toxins (*Chapter 2*) and substances that damage and kill cells and irritate the surface, causing the characteristic cough. Effective vaccines are available against both organisms, although 24 million infections of whooping cough occur annually worldwide. In contrast, *S. pneumoniae*, *H. influenzae*, *S. aureus* and *Moraxella catarrhalis* make up to 60% of the normal bacterial population of the nasopharyngeal mucous membrane in healthy individuals. They can become opportunistic pathogens in immunosuppressed individuals or following changes to the bacterium that render it increasingly virulent. The most common form of bacterial pneumonia is lobar pneumonia caused by *S. pneumoniae*, which results in a massive inflammation of one lobe of the lung. *S. aureus* may cause a bronchopneumonia.

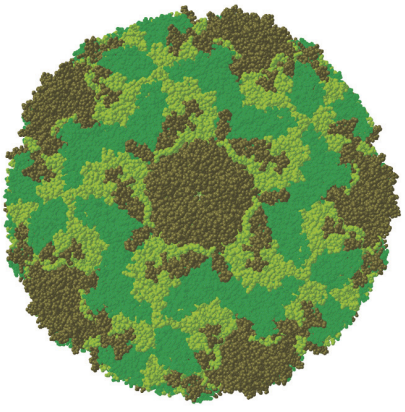


Figure 3.10 Molecular model of the capsid of a rhinovirus. PDB file 1C8M. ⏎

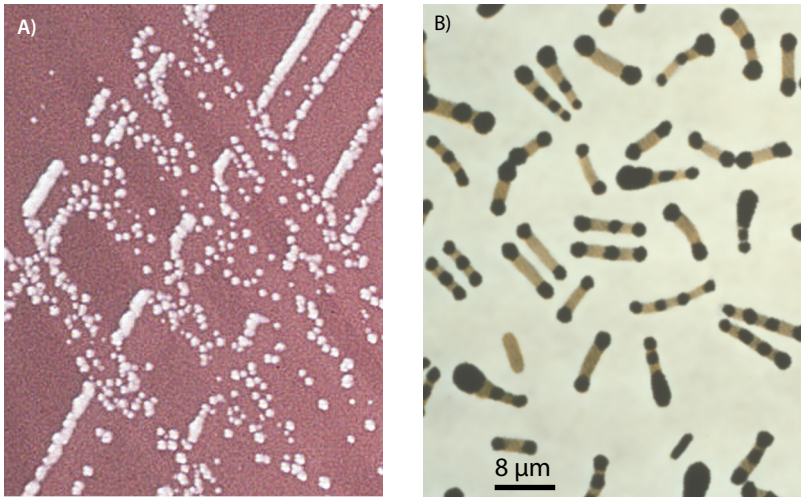


Figure 3.11 (A) The characteristic small metallic appearance of colonies of *Bordetella pertussis* growing on agar enriched with potato starch, glycerol and blood. Charcoal has been added to this medium to absorb bacterial metabolites that would inhibit the growth of the pathogen. (B) Photomicrograph of *Corynebacterium diphtheriae.* (A), Courtesy of the School of Biochemistry and Microbiology, University of Leeds, UK; (B), courtesy of the Centers for Disease Control and Prevention. ⏎

Infection with *Mycobacterium tuberculosis* (*Figure 3.12*), as suggested by the name, can develop into tuberculosis (TB) of the lungs and may be considered a rather special case of bacterial infection of the lower respiratory tract. The bacteria enter the alveoli in inhaled air and are phagocytosed by macrophages where they escape being killed and multiply (*Chapters 4* and *16*). Mycobacteria can then enter the lymphatic system and invade a neighbouring lymph node. The healing of local lesions leads to calcification of lung tissues. In immunodeficient individuals, the lymph nodes and tissues may be progressively affected until eventually the mycobacteria are spread in the circulation (*Chapters 13* and *14*). Also with impaired immunity, dormant *M. tuberculosis* can be reactivated, causing a severe form of pneumonia. Atypical pneumonias can result from infections with *Mycoplasma pneumoniae*, *Chlamydia pneumoniae* and *Legionella pneumophila*. These infections are associated with 'flu-like' symptoms, such as high temperatures and coughing, although bronchial secretions and sputum do not contain pus as would be expected of a typical bacterial lobar pneumonia.

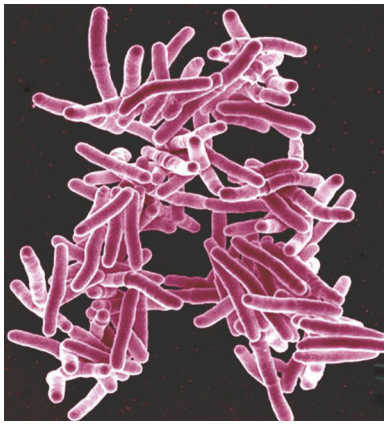


Figure 3.12 False colour electron micrograph of *Mycobacterium tuberculosis*. Courtesy of the National Institute of Allergy and Infectious Diseases. ⏎

Generally, pathogenic fungi do not produce toxins but damage tissues directly or disturb normal metabolic functions and can induce hypersensitivity responses (*Chapter 5*). Fungi can cause respiratory infections; *Aspergillus fumigatus* can invade the respiratory system and lead to one of several types of diseases. It may simply grow in the mucus of the bronchi and induce a hypersensitive state, but it can also invade old wound cavities of the lungs, such as those resulting from TB, and grow as a solid mass called an **aspergilloma**. Aspergillosis may also result from an invasive growth in the lungs and other tissues. Generally, the infective dose of spores is extremely large, although the invasive form may be secondary to other systemic diseases. Similarly, *Pneumocystis carinii* can cause a serious pneumonia in patients who are immunocompromised by acquired immunodeficiency syndrome (AIDS) (see *Box 3.3*). The yeast *Candida albicans* is also an opportunistic agent in sufferers of AIDS.

3.5 INFECTIONS OF THE GASTROINTESTINAL TRACT

All regions of the gastrointestinal tract (GIT) are subject to infections (*Figure 3.13*), although saliva traps and removes many pathogens and the acidic environment and proteolytic enzymes of the stomach kill most ingested microorganisms. Unfortunately, new ones are constantly introduced through breathing and eating.

Oral infections

Some bacteria can resist removal by saliva and become immobilised by binding to surface receptors of cells in the mouth, eventually forming biofilms and microcolonies. Oral streptococci, such as *Streptococcus sanguis* and *Streptococcus mutans* (*Figure 3.14*), secrete glycosyltransferases that mediate their adhesion to extracellular carbohydrates on tooth surfaces leading to the formation of dental plaque, which is a complex mixture

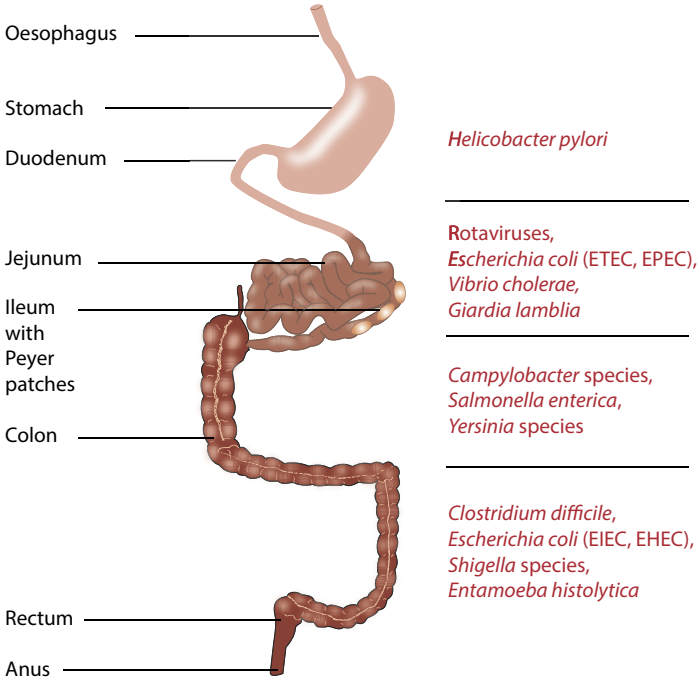


Figure 3.13 Some infectious organisms associated with the gastrointestinal tract. Enterohaemorrhagic (EHEC), enteroinvasive (EIEC), enteropathogenic (EPEC) and enterotoxigenic (ETEC) are different pathogenic forms of *Escherichia coli*.

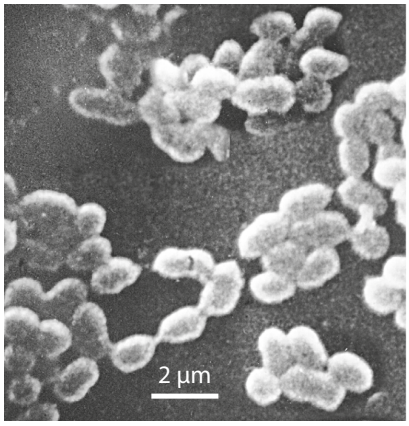


Figure 3.14 Electron micrograph of *Streptococcus mutans*. Courtesy of Professor J. Verran, School of Biology, Chemistry and Health Science, Manchester Metropolitan University, UK.

of bacteria and extracellular materials (see *Box 11.2*). These bacteria, together with *Actinomyces* species, can cause caries by forming plaque on the tooth enamel, where they catabolise sugars, producing acid that demineralises enamel and allows the dentine to be eroded. Abscesses of the roots of teeth can also be caused by mixed bacterial infections.

Periodontal (gum) diseases are inflammatory conditions that attack the gums, bone and other supporting structures of the teeth (see *Figure 11.8*). The extent of the inflammatory response depends upon the types of pathogens involved and the effectiveness of the immune response. However, they are major causes of tooth loss in adults. Gingivitis is the earliest form of periodontal disease and occurs when plaque accumulates on the teeth near the gums, which become inflamed and bleed easily. If detected and treated early, gingival tissues will return to normal without long-lasting damage. Untreated gingivitis progresses to periodontitis, which is also known as pyorrhoea. Plaque hardens and extends from the gum line to the tooth root, causing the gums to detach from the teeth and form pockets. Periodontal pockets create room for greater bacterial activity, particularly of facultative and obligate anaerobic bacteria, leading to a progressive cycle of tissue damage until eventually the bone supporting the teeth is destroyed, resulting in their loss.

Inflammation of oral tissues caused by the fungal infection **actinomycosis** often occurs following injuries, such as an accidental bite to the lining of the mouth or fracture of the jaw. Immunosuppression resulting from viral infections, HIV–AIDS (see *Box 3.3*), cancer treatment (*Chapter 17*) or treatment with broad-spectrum antibiotics can all allow the yeast *C. albicans* (*Figure 3.15*) to invade and colonise mucous membranes, eventually producing a thick layer of yeast cells called candidiasis or thrush.

Figure 3.15 Photomicrograph of *Candida albicans*. Courtesy of the Centers for Disease Control and Prevention/Dr Hardin.

Stomach and intestinal infections

Stomach and intestinal infections are caused by viruses, bacteria, protozoa and worms, all of which may be transmitted in contaminated food and/or drinking water or by faecal–oral contact. The need for strict personal hygiene is emphasised because these are the most frequent infections of children under five years of age. Approximately 40% of cases of diarrhoea (*Chapter 11*) in children are caused by rotaviruses (see *Figure 2.8*). In the very young, this is potentially lethal, and the CDC has estimated that diarrhoea diseases cause one in nine childhood deaths.

The motile bacterium *Helicobacter pylori* (see *Figure 11.39*) has specific receptors that enable it to bind to the gastric epithelium. It also secretes a urease, which catalyses the hydrolysis of urea, releasing ammonia that neutralises stomach acid and cytotoxins that damage the cells. This causes chronic inflammation of the gastric mucosal membrane and can lead to stomach and duodenal ulcers (*Chapter 11*). In the ileum and jejunum, nonenveloped viruses, such as adenoviruses and rotaviruses (see *Figures 1.2* and *2.8*), may infect enterocytes (*Chapter 11*) and damage the intestinal mucous membrane, with disruption of water and electrolyte resorption. This can result in intestinal cramps, vomiting, watery diarrhoea and an increased temperature. Enteropathogens, such as *Vibrio cholerae* and enterotoxigenic (ETEC) or enteropathogenic (EPEC) strains of *Escherichia coli* all have similar effects.

Peyer patches (see *Figure 3.13*) are areas of lymphoid tissue in the lower ileum composed of so-called microfold (M) cells, rather than the usual enterocytes and goblet cells (*Chapter 11*). These cells can translocate materials directly to the lymph follicles found beneath the mucosal surface. Invasive bacteria such as *Campylobacter jejuni* (*Figure 3.16*), Salmonellae and Yersiniae can use M cells to enter the submucosal area. Here they can multiply and destroy the adjacent epithelium, form abscesses and spread through the lymph and blood systems into the mesenteric lymph nodes, spleen and liver. The infection can also spread into the colon, causing inflammation of the colon or **colitis**. The ileum and colon can also be attacked by the bacteria *Yersinia enterocolitica*, *Salmonella enterica* and *C. jejuni*, resulting in abdominal cramps, vomiting, watery and occasionally bloody diarrhoea, and fever. *Shigella dysenteriae*, enterohaemorrhagic (EHEC) and EPEC *E. coli* pathotypes can cause a haemorrhagic colitis with bloody stools and subfebrile–febrile temperatures. *Clostridium difficile*, a normal inhabitant of the gut, is an opportunistic pathogen. It is especially common in older people in hospitals and nursing homes and has been implicated in **iatrogenic** infections following medical interventions, such as antibiotic therapy. Infection with *C. difficile* is recognised as the major causative agent of colitis and diarrhoea, which may occur following antibiotic intake and can be fatal in elderly patients.

The waterborne protozoan parasites *Giardia lamblia* (*Figure 3.17*) and *Cryptosporidium parvum* can infect the GIT. *G. lamblia* frequently causes chronic disease, with watery diarrhoea and, in some cases, a subfebrile temperature leading to malnutrition in children because of malabsorption. *C. parvum* can adhere to the epithelium of the small intestine and cause a shortening of the villi, which causes diarrhoea. The pathogenic protozoan *Entamoeba histolytica* is thought to infect 50 million people and kill about 100 000 per year globally due to amoebic liver abscesses.

The GIT is subject to attack by a variety of helminth parasites, including types of roundworms and tapeworms (*Section 3.11*). Abdominal pain and diarrhoea are frequent symptoms.

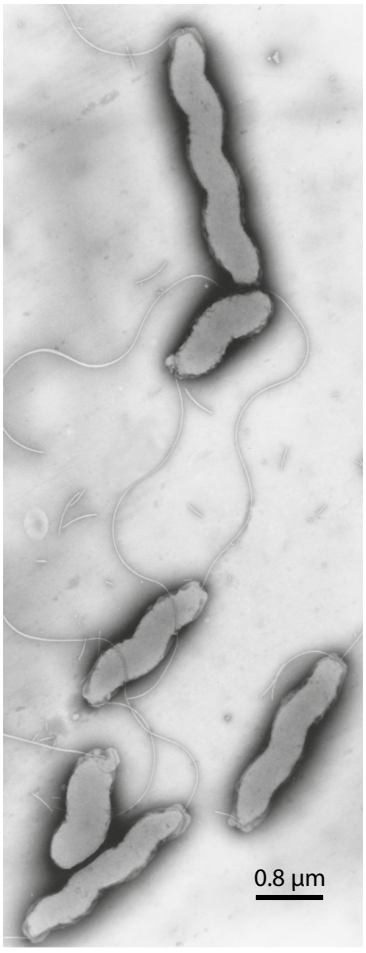


Figure 3.16 Electron micrograph of *Campylobacter jejuni*. Courtesy of Dr A. Curry, Manchester Royal Infirmary, UK. ⏎

3.6 INFECTIONS OF THE UROGENITAL SYSTEM

The urinary and genital systems are subject to infections by specific pathogens (*Figure 3.18*). Urinary tract infections (UTIs) are frequent in the developed world, with many millions of cases occurring annually. Several factors, including diabetes mellitus (*Chapter 7*), scarring, kidney stones, use of catheters or anatomical peculiarities of the urinary tract, all predispose individuals to UTIs. *E. coli* is the most common agent, causing up to 80% of the relatively uncomplicated UTIs, such as cystitis – a short-lived, acute infection. Cystitis originates in the perianal area and moves up the urethra into the bladder. Women are more prone than males to cystitis because of their relatively short urethra and the hormonal changes associated with the menstrual cycle (*Chapter 7*). Cystitis is readily treated by increasing the intake of fluids, particularly of acid drinks like cranberry juice, which causes increased flushing of the bladder, and using antibiotics. Unfortunately, reinfections are frequent.

P. aeruginosa (see *Figure 2.3*) has been known to cause UTIs following hospitalisation. Similarly, infections by *Enterococcus faecalis* and *Klebsiella pneumoniae* have followed organ transplants (*Chapter 6*). Unfortunately, these organisms often show multidrug resistance (MDR) against antibiotics (*Box 3.2*; see also *Section 3.11*). Thrush in the bladder, caused by the yeast *C. albicans*, can also occur following antibiotic treatment.

There has been a large increase in the incidence of sexually transmitted diseases (STDs). For example, in England in 2023, there were 401 800 cases of STDs, an increase of 4.7% compared with the 383 789 in 2022. Sexually transmitted diseases can affect the urinary system and the genital tract and are caused by a wide range of pathogens. Human papilloma virus (HPV) can cause anogenital warts (condyloma acuminata).

Figure 3.17 Electron micrograph showing the ventral view of *Giardia lamblia*. Courtesy of David Gregory and Debbie Marshall, published under CC BY 4.0 licence. ⏎

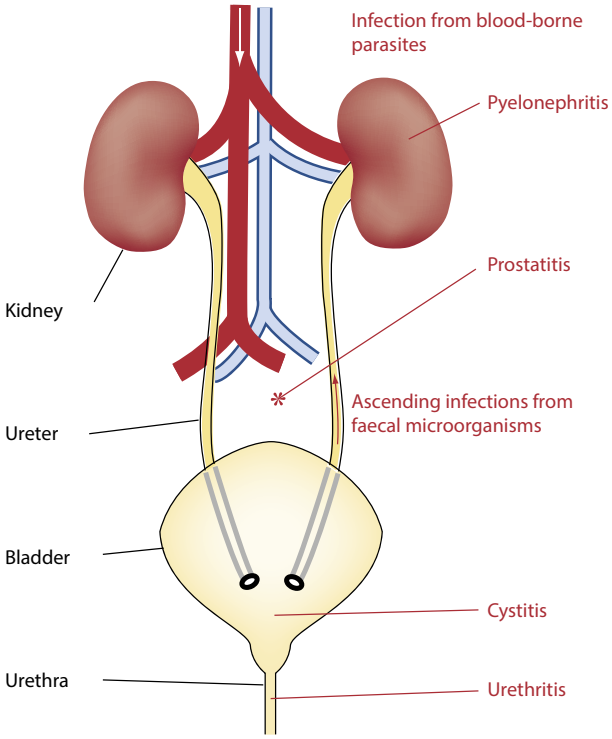


Figure 3.18 Examples of infections associated with the urogenital system. ⏎

BOX 3.2 DEVELOPMENT OF DRUG RESISTANCE ⏎

Primary resistance describes a pathogen or parasite that is not initially susceptible to a particular drug. However, *secondary* or *acquired* resistance develops over time in a previously sensitive organism. *Phenotypic* resistance describes the situation where a subpopulation of the organism is able to grow in concentrations of the drug that would normally inhibit its reproduction. *Tolerance* is characterised by a reduced rate of pathogen death and is an additional cause of drug resistance.

Unfortunately, as emphasised throughout *Section 3.11*, secondary resistance to drugs has evolved in all groups of pathogens and parasites. Antibiotic resistance is achieved by four major biochemical mechanisms. Firstly, mechanisms can evolve that decrease the intracellular concentration of the drug by reducing its entry into or increasing its expulsion from the cell. Secondly, the organism may activate enzymes, or new ones may evolve, that are able to modify or inactivate the drug. Thirdly, when an antibiotic blocks a single reaction, for example by inhibiting an enzyme, then resistance may be acquired by the acquisition of alternative mechanisms that bypass the reaction. Fourthly, the ability to modify biochemically the targets of specific drugs can develop, ensuring the target is no longer recognised. These methods are not exclusive, and organisms may use more than one way of nullifying the effects of a single drug.

In viruses, resistance normally results from spontaneous mutations in viral genes that result in structural changes to the target so that it is no longer recognised by the drug. Resistance against a particular drug also develops in previously sensitive bacteria because of spontaneous mutations. It may, however, occur by the acquisition of new genes from other bacteria. This genetic material may be chromosomal or extrachromosomal. Chromosomal material most commonly consists of transposons. Extrachromosomal genetic information is usually contained in plasmids, which are particularly prominent in transferring antibiotic resistance between resistant and nonresistant bacteria. Further, a single plasmid may contain the genetic determinants for resistance to several drugs. Antibiotic resistance develops in eukaryotic pathogens: protozoa, fungi and helminths. Resistance in the malarial parasite is described in *Chapter 2*. Sadly, strains of all types of pathogens and parasites now display multidrug resistance and are often referred to as 'superbugs', for example methicillin-resistant *Staphylococcus aureus* (*Figure 3.19*). Indeed, in such cases, clinicians have to resort to using drugs of last resort, such as vancomycin and linezolid.

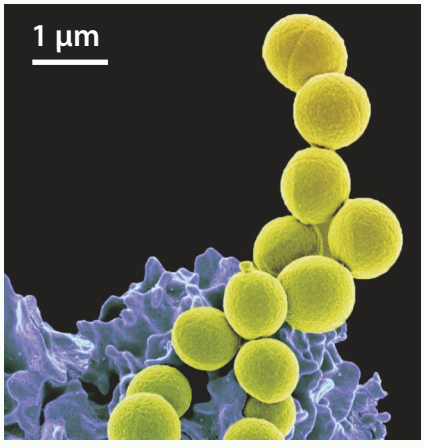


Figure 3.19 False colour electron micrograph of methicillin-resistant *Staphylococcus aureus* being ingested by a neutrophil. Courtesy of the National Institute of Allergy and Infectious Diseases. ⏎

In all cases, the continual emergence of resistant strains of pathogens and parasites means there is a need to develop new pharmacological agents capable of treating infectious diseases. Unfortunately, the development of new drugs has woefully lagged behind the evolution of resistance. Several reasons have been advanced for this, including poor incentives for pharmaceutical companies to invest in antibiotic development because of poor successes, particularly in identifying novel 'lead' compounds on which to establish families of new drugs; limited markets because of the large number of existing antibiotics; and because most customers (patients!) generally only use the antibiotic for one bout of illness.

Strains 16 and 18 of HPV cause lesions in the cervix that are involved in the development of cervical carcinoma (see *Figures 17.24–17.26*).

Infection with HHV-2 (see *Table 3.1*) can lead to genital herpes with painful ulcers and vesicular lesions of the genital mucous membrane. Infection with HIV leads to AIDS (*Box 3.3*).

The most widely known STDs are probably gonorrhoea and syphilis, caused by the bacteria *Neisseria gonorrhoeae*, a Gram-negative diplococcus, and *Treponema pallidum*, a spirochete, respectively (*Figure 3.25*). Gonorrhoea, a common STD, is a pelvic inflammatory disease whose major symptoms include a purulent inflammation of the uterine cervix and urethritis. In some women, however, the infection may be relatively

BOX 3.3 HIV AND AIDS

In March 1981, physicians in New York reported eight cases of the previously rare Kaposi sarcoma (KS) in a form that was markedly more aggressive than usual. All those affected were young gay men. Also in 1981, the Centers for Disease Control and Prevention (CDC) in Atlanta, Georgia, began to investigate an increase in requests from New York and Los Angeles for pentamidine (*Section 3.11*), a drug used to treat *Pneumocystis carinii* pneumonia (PCP), an extremely serious fungal infection. In addition to PCP and/or KS, patients suffered a variety of opportunistic pathogens that eventually caused their death. The combination of infections and KS appeared to be the result of a total breakdown of the immune system and became known as acquired immunodeficiency syndrome (AIDS). Subsequently, AIDS was seen in intravenous drug abusers and in a number of recipients of blood transfusions, and a viral cause was suspected. In 1983, a virus, variously named lymphadenopathy virus (LAV) and human T cell lymphotropic virus type III (HTLV III), was isolated. Following international agreement, the virus was renamed human immunodeficiency virus (HIV; *Figure 3.20*) in 1986. Several strains of the virus have since emerged; HIV-1 and -2 are the most prevalent, with the former being the more destructive. The World Health Organization (WHO) estimates that in the 40-or-so years since its emergence, the HIV–AIDS viruses have caused a pandemic, infecting approximately 76 million sufferers and causing about 22 million deaths. In 2020, it estimated that 38 million patients were infected, with roughly 1.5 million new cases occurring and nearly 700 000 deaths, of which nearly 100 000 were children.

A simplified replication cycle of HIV is shown in *Figure 3.21*. Glycoproteins in the envelope of the virus bind to CD4 proteins on target cells. The subsequent fusion of the viral envelope and host cell plasma membrane facilitates entry of the capsid into the host cell. Uncoating of the viral genome initiates the first stage of viral replication. Human immunodeficiency virus is a retrovirus, that is, its genome is RNA. Its reverse transcriptase (RT) catalyses three types of reactions. Firstly, it synthesises a complementary negative-strand DNA copy of the RNA genome (see *Table 2.3*). Secondly, it hydrolyses the original RNA before, thirdly, forming a complementary positive strand to give a double-helical DNA viral genome (*Figure 3.22*). The DNA enters the nucleus, where it is spliced into host chromosomal DNA in a reaction catalysed by the HIV integrase. Transcription of the viral genes by the host RNA polymerase II produces viral genomic RNA and viral messenger RNA. Translation of the latter by host ribosomes produces polyproteins that are hydrolysed by an HIV protease to give native viral proteins. Viral glycoproteins are formed by the endomembrane system of the cell and transferred into the plasma membrane. Here, virions are assembled on its inner surface and bud from its surface, acquiring their envelopes as they do so.

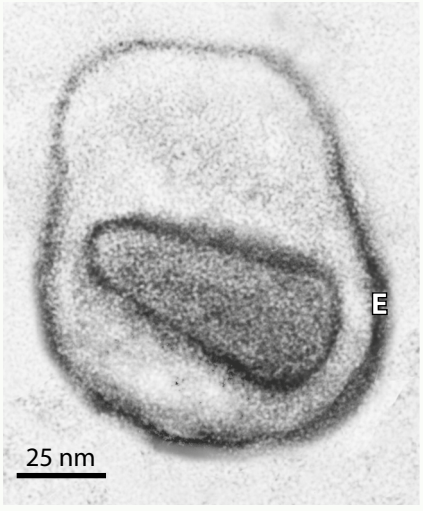


Figure 3.20 Electron micrograph of a single human immunodeficiency virus (HIV) virion. E denotes its envelope.

Given that HIV infects CD4+ helper T lymphocytes, which are key regulatory cells of the immune system (*Chapter 4*), the patient is rendered susceptible to a whole range of disease from all infectious microbial groups (*Table 3.2*) and a variety of tumours (*Table 3.3*).

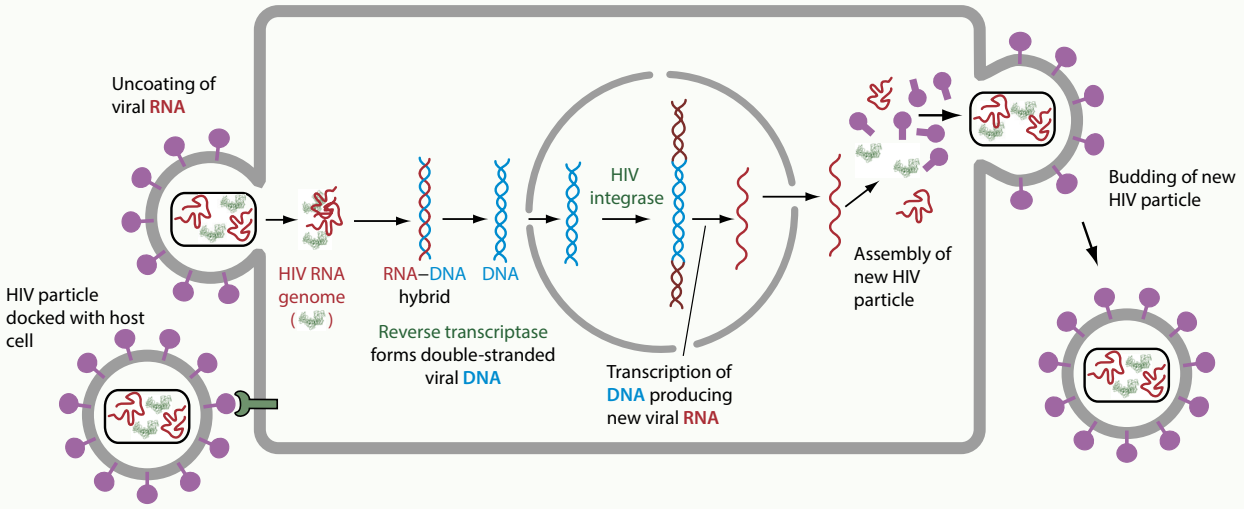

Uncoating of viral RNA
HIV particle docked with host cell
HIV RNA genome
RNA–DNA hybrid
DNA
Reverse transcriptase forms double-stranded viral DNA
HIV integrase
Transcription of DNA producing new viral RNA
Assembly of new HIV particle
Budding of new HIV particle

Figure 3.21 A simplified illustration of the human immunodeficiency virus (HIV) replication cycle.

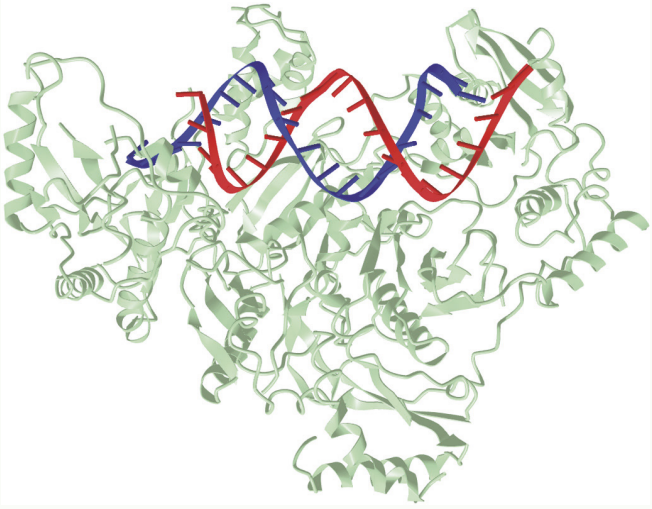


Figure 3.22 Molecular modal of the human immunodeficiency virus 1 (HIV-1) reverse transcriptase with a bound RNA (red)–DNA (blue) hybrid duplex. PDB file 4B3P.

Disease	Cause	Infectious agent
Cryptococcosis	*Cryptococcus neoformans*	Fungus
Diarrhoea and dysentery	*Cryptosporidium parvum*	Protozoan
Kaposi sarcoma	Kaposi sarcoma virus (HHV-8)	Virus
Lymphoma	Epstein–Barr virus (HHV-4)	Virus
Mucocutaneous thrush	*Candida albicans*	Yeast
Pneumonia	*Pneumocystis carinii*	Fungus
Shingles	*Varicella zoster* (HHV-3)	Virus
Tuberculosis	*Mycobacterium tuberculosis*, *Mycobacterium avium*	Bacteria

HHV, human herpes virus.

Table 3.2 Some diseases associated with human immunodeficiency virus (HIV)–acquired immunodeficiency syndrome (AIDS).

Types of cancers
Anal cancer
Cervical cancer[a]
Liver cancer
Hodgkin lymphoma
Kaposi lymphoma[a]
Non-Hodgkin lymphomas[a]

[a] These are considered human immunodeficiency virus (HIV)–AIDS-defining cancers when combined with a positive result for HIV antibodies in the blood and a CD4+/CD8+ T lymphocyte ratio of ≤0.25 (reference value 2.2).

Table 3.3 Examples of virus-related cancers to which acquired immunodeficiency syndrome (AIDS) patients are increasingly susceptible.

The protozoan *Cryptosporidium parvum* can cause moderate to severe diarrhoea that would soon be resolved in healthy people. In HIV-infected patients, cryptosporidial diarrhoea is among the most common clinical presentations during the transition to full-blown AIDS, particularly in developed countries. The diarrhoea is severe and protracted and may become life threatening. Death is commonly caused by pneumonia associated with the fungus PCP among a host of opportunistic infections.

Extensive educational programmes have emphasised some of the dangers of unprotected casual sex with different partners and the sharing of needles among intravenous drug addicts. There is no cure for AIDS and, to date, despite extensive efforts using multiple approaches, a successful vaccine is still not available, although a vaccine designated HVTN706/MOSAICA is currently undergoing preclinical testing on transgender men in the USA and Europe. However, drugs have been developed to treat HIV that reduce the blood viral load to low or even nondetectable levels (*Table 3.4*) by acting at different stages in the viral replication cycle. Some, like enfuvirtide and maraviroc, interfere with the binding of viral glycoproteins to the host cell plasma membrane. Nucleoside RT inhibitors are prodrugs that are converted to their triphosphate derivatives

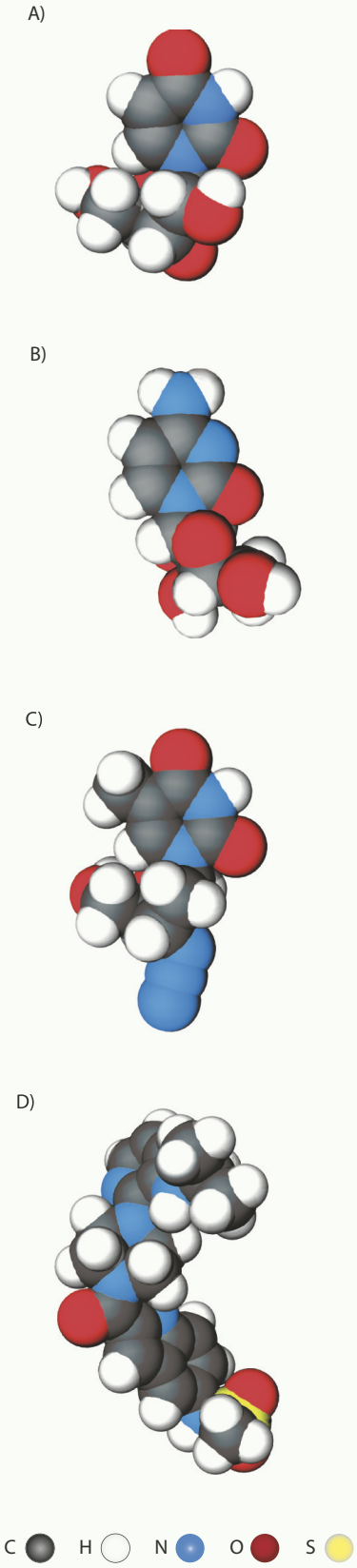


Figure 3.23 Computer drawn models of the ribonucleosides (A) uridine, (B) cytidine and (C) the anti-human immunodeficiency virus (HIV) drug nucleoside reverse transcription (RT) inhibitor azidothymidine (AZT). Note their structural resemblance to that of (D), the non-nucleoside RT inhibitor delavirdine.

Drug inhibits	Examples[a]
Binding/fusion	Enfuvirtide, maraviroc
Reverse transcriptase (RT)	Nucleoside RT inhibitors: azidothymidine (AZT), dideoxyinosine, tenofovir Non-nucleoside RT inhibitors: delavirdine, efavirenz, rilpivirine
Integrase	Raltegravir
Protease	Amprenavir, darunavir, ritonavir

[a] Not comprehensive; approximately 30 anti-HIV drugs are currently in users.

Table 3.4 Examples of drugs used to treat human immunodeficiency virus (HIV).

by cellular kinases. These are structural analogues of substrate for the enzyme but their incorporation into the growing nucleotide strand leads to chain termination. Non-nucleoside RT inhibitors, such as delavirdine (*Figure 3.23*), also inhibit RT activity, but by binding to it at allosteric positions rather than to its active site. The integrase inhibitor raltegravir prevents the ligation of viral and host DNA molecules. Protease inhibitors are peptide analogues of its substrate that bind to the active site (*Figure 3.24*). They are therefore competitive inhibitors of the enzyme and prevent maturation of the viral polypeptides.

Figure 3.24 Molecular model of the human immunodeficiency virus 1 (HIV-1) protease complexed with its inhibitor, amprenavir, shown in red. See also *Table 3.4*. PDB file 3NU3.

Patients with AIDS at any stage are treated by highly active antiretroviral therapy (HAART; see *Margin Note 16.1*), which typically consists of a nucleoside RT inhibitor and either a non-nucleoside RT inhibitor or a protease inhibitor (see *Table 3.4*). Over 30 different anti-HIV drugs are available, meaning that many combinations can be used and, indeed, tailored to the needs of the individual. Infected people can now expect to live relatively healthy and fulfilling lives following their initial infection. However, combination therapy must be taken for the remainder of the patient's life. Also, the aggressive therapy can lead to problems, given that the drugs have numerous side effects. Multiple drug-resistant strains of HIV have now evolved, which emphasises the need for constant research into new therapies.

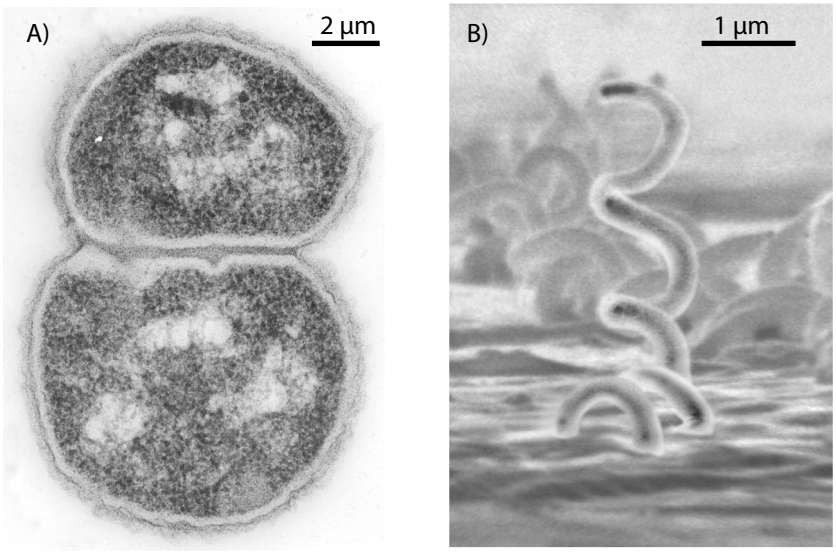


Figure 3.25 Electron micrographs of (A) *Neisseria gonorrhoeae* and (B) *Treponema pallidum*. (A), Courtesy of Dr A. Curry, Manchester Royal Infirmary, UK; (B), courtesy of the Centers for Disease Control and Prevention. ⏎

asymptomatic and may go unnoticed. In men, symptoms include a discharge from the tip of the penis that resembles pus in appearance, pain during micturition and possibly a painful swelling in one testicle.

Syphilis was thought to have originated in the Americas and been brought to Europe by sailors on the Columbian expeditions. Other evidence suggests that it was present in the 'Old World' long before these events. *Treponema* spirochetes can enter through mucous membranes or minute abrasions in the skin during sexual acts. The infection shows three stages of pathogenesis. Initially, an ulcerous **chancre** develops at the site of infection. The infection then spreads to nearby lymph nodes, causing them to swell and harden. Secondary syphilis develops after one to three months. It is characterised by the presence of highly infectious lesions on various parts of the body. The disease may lie dormant for many years but, if not treated with antibiotics, will develop into tertiary syphilis causing inflammations of the aorta and CNS. Dementia, heart attacks and death can all result. Patients with tertiary syphilis cannot infect others with the disease. Some patients may develop benign late syphilis, which is usually rapid in onset but does respond well to treatment. It usually begins 3–10 years after infection and is characterised by the development of **gummas**. These are tumour-like growths of a rubbery consistency that are most likely to affect the skin or long bones, but can also develop in the eyes, mucous membranes, throat, liver and stomach lining. However, since the use of antibiotic treatments for syphilis, they have become relatively less common.

The protozoan *Trichomonas vaginalis* (*Figure 3.26*) is a frequent coloniser of the mucosal membrane of the urogenital system. It is generally asymptomatic, but an inflammatory reaction, trichomoniasis vaginitis, may result in a frothy cream-coloured discharge. Some types of *C. trachomatis* (*Figure 3.7*) are among the most common causes of STDs, producing local inflammations of the urethra and cervix. In contrast, other types are highly invasive and infiltrate the lymphatic system leading to necrosis (lymphogranuloma venereum).

Several arthropods may be sexually transmitted. These include pubic lice and the scabies mite (see *Figure 2.45*).

3.7 SEPSIS AND SYSTEMIC INFECTIONS

A relatively small number of pathogens are able to enter the body in the lymphatic or circulatory systems and produce a generalised or systemic infection involving numerous body organs: brain, bone marrow, kidneys, liver, lungs and spleen. Local skin infections, such as with *S. pyogenes* or *S. aureus*, or infections of the GIT or urogenital system with Enterobacteriaceae can, unfortunately, progress to acute generalised infections within only a few days, given appropriate conditions. For example, *S. pyogenes* or *S. aureus* secrete **superantigen** toxins (*Chapters 2* and *4*) that stimulate the release of cytokines

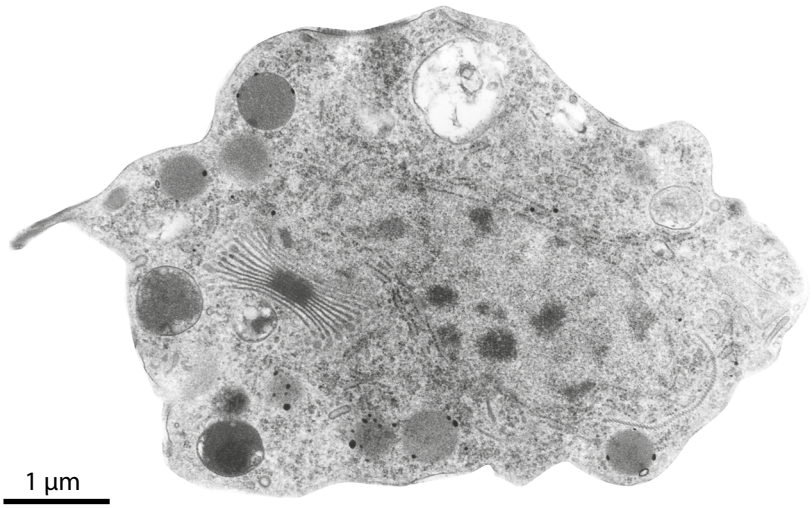


Figure 3.26 Electron micrograph of *Trichomonas vaginalis*. Courtesy of Dr A. Curry, Manchester Royal Infirmary, UK. ⏎

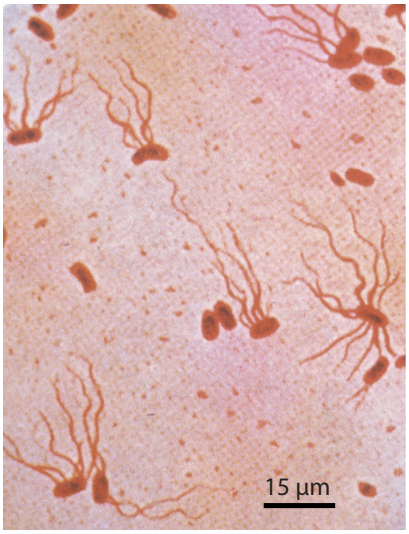


Figure 3.27 Photomicrograph of *Salmonella typhi*. Courtesy of the Centers for Disease Control and Prevention. ⏎

from immune cells and produce an excessive inflammatory response called **systemic inflammatory response syndrome** (SIRS). This consists of four discrete stages.

1. Stage I is characterised by increases in respiration and heart rates and body temperature; the leukocyte count may be increased or decreased.
2. Stage II is **sepsis**, the presence of the microorganism in the blood of the patient. Normally, less than 50% of cases can be identified. This stage may be cured intrinsically by the immune system or extrinsically by using antibiotics but, if this is unsuccessful, the condition progresses to the third stage.
3. Stage III is denoted by serious sepsis or **multiorgan dysfunction syndrome** (MODS), characterised by lactic acidosis, decreased blood pressure, hypoxia and oliguria. If clinical measures fail to stabilise the patient, it then progresses to the final stage.
4. Stage IV consists of septic shock with irreversible organ failure and the death of the patient, in most cases.

In contrast with the possible acute developments of localised infections, generalised infections may have incubation times of up to three weeks. A classic example of such a systemic infection is typhoid fever caused by *Salmonella typhi* (*Figure 3.27*). The pathogen may be ingested and can enter the body through the tonsils in the throat and Peyer patches in the GIT. The bacterium is distributed to various organs in the lymph and blood, infects cells and grows intracellularly. The major symptoms are fevers of up to 40°C and the development of abscesses in the intestine that may cause perforations and fatal peritonitis. Even when the patient recovers, a relapse may occur. *S. typhi* infects 11–21 million patients annually, mainly in the developing world. If untreated, it can be fatal. Some recovered patients retain the pathogen in the gallbladder and excrete bacteria in their faeces, becoming carriers, although remaining healthy themselves.

The systemic condition Lyme disease, also called Lyme arthritis or Lyme borreliosis, is caused by the spirochete bacterium *Borrelia burgdorferi* (see *Figure 2.44*). It was first detected in 1975 in Lyme, Connecticut, USA. *B. burgdorferi* is spread by blood-sucking ticks of the genus *Ixodes* (see *Figure 2.45*) that become infected when feeding on animal reservoirs of the bacterium. Lyme disease is the most prevalent arthropod vector-borne disease in the northern hemisphere, affecting an estimated 476 000 people in the USA and more than 200 000 in western Europe annually.

The disease presents in three stages:

1. early localised
2. early disseminated
3. late disseminated

The early localised stage is a red rash or erythema migrans that appears at the site of the tick bite approximately seven days later. Interestingly, the rash may not be itchy or painful. Other early symptoms may include chills, fever, headaches and tiredness. The early disseminated infection occurs within days or weeks following the early local

Margin Note 3.2
Salmonella typhimurium

Salmonella typhimurium (*Figure 3.28*), as its name suggests, causes a lethal, systemic infection in mice that resembles typhoid fever. However, in humans, it does not cause as severe a disease as *S. typhi*, although it is a major cause of food poisoning. This rarely leads to fatalities, other than in the elderly or very young or in patients with depressed immune systems who are not treated with antibiotics. The disease is characterised by symptoms of nausea and vomiting, abdominal cramps and diarrhoea that generally last up to seven days.

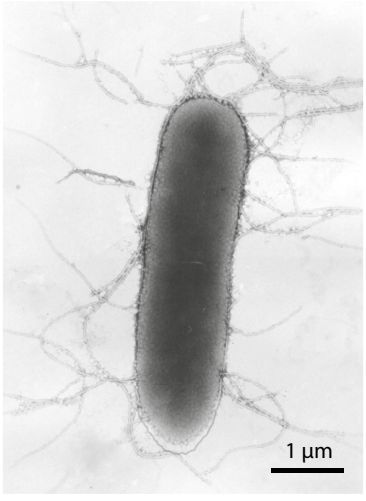


Figure 3.28 Electron micrograph of *Salmonella typhimurium*. Courtesy of Dr A. Curry, Manchester Royal Infirmary, UK. ⏎

stage as the lymphatic and cardiovascular systems transport the *Borrelia* bacteria throughout the body. Transient joint and muscle pains are common at this point, and, in up to a fifth of patients, erythema migrans rashes develop at sites other than that of the tick bite. Late disseminated infection occurs after several months. It presents as chronic symptoms that affect many areas of the body, including the brain, eyes, heart, joints and nerves. Up to 60% of untreated patients develop Lyme arthritis, typically starting about six months after infection, and 5% may experience chronic neurological symptoms. Even if appropriately treated, 10–20% of patients develop joint pains, memory problems and tiredness for at least six months. Repeated episodes of joint pains and swellings may occur months to years later in untreated patients.

Diagnosing early localised Lyme disease is difficult. Molecular biological-based diagnosis techniques, such as those based on the polymerase chain reaction (PCR; *Section 3.10*) or enzyme-linked immunosorbent assay (ELISA; *Chapters 4* and *17*), are necessary to investigate the later stages. The disease is treatable with antibiotics, especially doxycycline but also amoxicillin or ceftriaxone; however, if detected and treated at a very late stage, permanent damage may have occurred to the joints, heart or nervous system.

The zoonosis brucellosis, caused by various species of *Brucella*, is also a primary systemic infection. These are common infections of cattle, sheep (*Brucella abortus*) and goats (*Brucella melitensis*) and can infect humans if they drink contaminated milk or inhale the bacteria into the lungs or if there is direct contact through skin lesions. People who work in direct contact with animals, such as farmers, veterinarians, shepherds, goatherds and abattoir workers, are most at risk. The pathogen colonises and grows in the main abdominal organs and bone marrow, producing periodic attacks of relapsing fever over an extended period.

3.8 PREVENTING INFECTIOUS DISEASES

Infections are the most common cause of human morbidity and mortality. In developing countries at the beginning of the last century, diseases (for example TB, pneumonia and bacterial infections secondary to influenza) were major causes of death. However, in the developed world, their impact has been lessened by public health measures, such as the provision of potable water supplies, efficient removal of sewage and other waste, improved housing, better sanitation and advanced social and economic conditions (*Chapter 2*). Unfortunately, in developing countries, diseases such as malaria, TB, respiratory and GIT infections, and HIV–AIDS are still major causes of death, particularly in children.

Preventative care includes health education provision to promote healthy lifestyles. The strict imposition of hand-washing protocols can have dramatic effects in reducing the spread of pathogens (*Figure 3.29*). The isolation of infected patients is sometimes an essential feature of preventative medicine, as are screening programmes and the widespread use of vaccines.

Vaccines and vaccination

Vaccines are biological agents that can elicit appropriate immune responses in an individual to a pathogen or parasite (*Figure 3.30*). They can be weakened, killed or fragmented microorganisms, toxins, proteins or nucleic acids. The simplest, though not necessarily the most effective, method of preparing

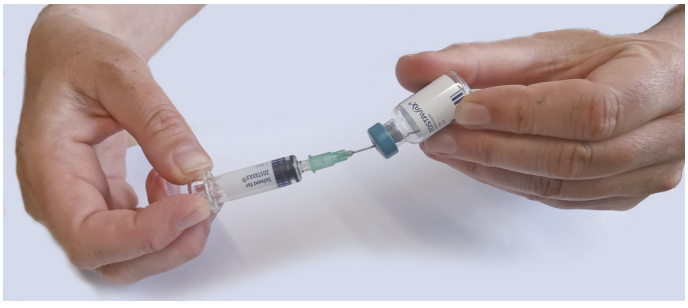


Figure 3.30 A general-practice nurse preparing a shingles vaccine for injection into a 70-year-old man as part of the NHS (UK) shingles vaccination programme. Courtesy of Fiona Sanders, Bracondale Medical Centre, Stockport, UK.

Step 1
Squeeze a small amount of sanitiser gel or soap over left palm. Dip all the fingers of the right hand into the left palm and vice versa
Step 2
Join hands palm to palm
Step 3
Rub right palm over left hand as shown and then left palm over right hand
Step 4
Interlace fingers with hands joined palm to palm
Step 5
With fingers still interlaced, rub the backs of fingers against opposing palms
Step 6
Clasp thumbs in opposing palms and wash with rotational movements
Step 7
Wash wrists with rotational movements of opposing hand
Rinse and dry thoroughly

Figure 3.29 Recommended stages in hand washing.

vaccines is to kill whole microorganisms. These are then injected into the host to induce an immune response. An example of a 'killed' vaccine is that used to protect against whooping cough caused by *B. pertussis.* In some instances, the immune response produced against dead organisms is insufficient to induce good immunity. This is usually because killing microorganisms often denatures their proteins, so, while the immune response recognises the denatured proteins, it does not react to the native proteins on the pathogen. To overcome this problem, live but attenuated (weakened) microorganisms may be used. These microorganisms are less virulent and, in most cases, stimulate an effective immune response in the host. This type of approach is used for the combined vaccine against measles, mumps and rubella (MMR) and for the oral vaccine against poliomyelitis (*Box 3.4*).

Some patients do not develop an effective immune response to these weakened microorganisms and, unfortunately, the weakened microorganisms can become virulent again, a phenomenon known as **reversion**. To overcome this problem, selected proteins from, for example, bacterial capsules or viral envelopes are extracted from the pathogens and used as vaccines. These are known as subunit vaccines. Subunit vaccines against viral proteins can be produced cheaply by employing recombinant DNA techniques. In such cases, DNA coding for the protein in question is isolated and cloned. This DNA is then transfected into a suitable microorganism that can be cultured and induced to biosynthesise the protein. An example is the recombinant vaccine against hepatitis B produced in the yeast *Saccharomyces cerevisiae* transfected with the gene encoding the S (surface) protein of the virus. Unfortunately, bacterial capsular polysaccharide is often poor at

BOX 3.4 BABY AND CHILDHOOD VACCINATION PROGRAMMES

Most countries of the developed world have baby and childhood vaccination programmes that protect against a variety of potentially life-threatening conditions. In the UK, the NHS vaccination schedule for babies and children spans from 8 weeks to 15 years of age. The 6-in-1 vaccine, which protects against diphtheria, hepatitis B, Hib (see *Margin Note 3.1*), polio, tetanus and whooping cough (pertussis) infections, is given to 8-week-old babies, with two booster doses at 12 and 16 weeks. Babies who are allergic to the vaccine, have a high temperature or suffer neurological problems, including poorly controlled epilepsy, are not eligible for the vaccine. Vaccines against rotavirus (see *Figure 2.8*) and meningococcal group B bacterial (MenB) infections also begin at eight weeks. The former is given as two doses, the second at 12 weeks. The MenB vaccine protects against meningococcal group B bacteria that cause meningitis and sepsis, which can lead to severe brain damage, amputations and, sometimes, death. The initial vaccination at 8 weeks is boosted by a second when the baby is 16 weeks old. Babies given the MenB vaccine are likely to develop a high temperature within 24 hours of vaccination that is treatable with paediatric-strength liquid paracetamol.

The NHS vaccination regime also involves two doses of pneumococcal vaccine, also known as the pneumonia vaccine, at 12 weeks and at 1 year of age. This protects against *Pneumococcus* bacteria that cause ear infections, pneumonia, sepsis and meningitis. One-year-old babies also receive a combined Hib/meningitis C vaccine, which is followed a month later with the combined measles, mumps and rubella (MMR) vaccine. A second dose of MMR vaccine is given at the age of three years and four months.

Since August 2022, an annually administered nasal spray influenza vaccine has been available for 2- to 11-year-old children. An injected vaccine can be offered to children if the nasal spray vaccine is unsuitable for them. Children of both sexes are offered the human papillomavirus (HPV) vaccine at 12–13 years, which helps protect against cancers caused by HPV. These include cervical cancer and some cancers of the anal, genital, and head and neck areas. Those aged 13–15 years are routinely offered, as part of the NHS vaccination programme, the 3-in-1 teenage and the MenACWY vaccines. The 3-in-1 vaccine boosts protection against tetanus, diphtheria and polio. The MenACWY vaccine protects against four strains of meningococcal bacteria: A, C, W and Y.

Unfortunately, approximately 1.5 million children under five years die annually from vaccine-preventable diseases globally, mainly because of the lack of access to essential childhood vaccines. Thus, the key objectives of the Immunization Agenda 2030 of the World Health Organization (WHO) are to provide access to these vaccines and reduce immunisation inequalities for all children.

Margin Note 3.3 Jenner and his vaccination legacy

Protecting people from infection by deliberately exposing them to components of pathogens or parasites was initiated in modern times by Edward Jenner in 1796. He showed that immunity to smallpox, caused by the variola virus, was induced by the introduction into the skin of material from the crusts of cowpox lesions. This process became known as **vaccination** (from the Latin *vacca*, 'cow'), a term still used today. In this case, immunisation with the cowpox virus induced immunity that cross-reacted with the variola virus, because they share some molecular similarities. Cross-reactions between vaccines are quite rare and most vaccines are usually specific for only one type of microorganism. The WHO used a cross-reacting vaccinia virus-based vaccine (*Figure 3.31*) to successfully eradicate smallpox as described in

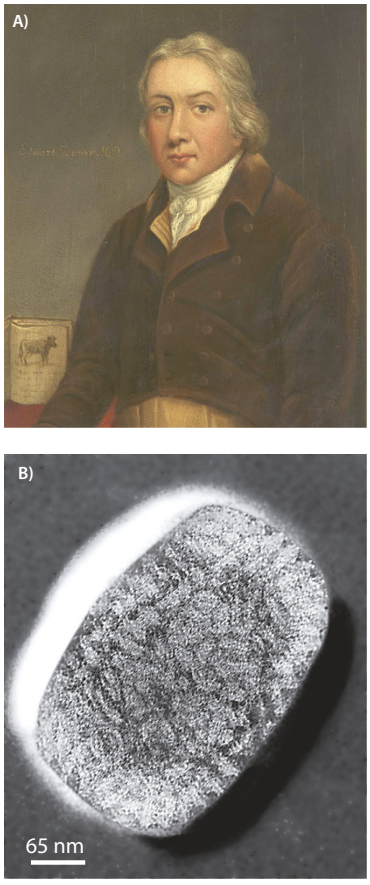


Figure 3.31 (A) Edward Jenner (1749–1823), the father of vaccination in Western medicine. (B) Electron micrograph of vaccinia virus. Courtesy of North West Regional Virus Laboratory, Booth Hall Hospital, Manchester, UK.

stimulating immunity. To render the vaccine more immunogenic, the polysaccharide is attached to an immunogenic protein. Such vaccines are known as **conjugate vaccines**: examples include vaccines against *N. meningitidis* serogroup C and *H. influenzae*.

For those infectious diseases where toxins (*Chapter 2*), rather than the microorganism, are responsible for the disease, vaccines may be prepared against chemically modified or heat-inactivated toxins. These inactive toxins, known as **toxoids**, are then used for vaccination purposes and, indeed, this is the approach used for diphtheria and tetanus vaccines.

Vector vaccines are produced by inserting genes for antigens of a pathogen into a nonpathogenic viral vector, for example an adenovirus. Examples of this type of vaccine include some of the Ebola, coronavirus disease 2019 (COVID-19) and smallpox vaccines. Since several strains of a pathogen may cause a disease, multivalent vector vaccines containing components from each may be used in such cases.

Gene-based vaccines are the newest development in vaccine production. The DNA or mRNA vaccines rely on an antigen-coding nucleic acid being taken up and expressed by cells of the recipient to induce an immune response. In vaccines using DNA, this must first be transcribed to form the corresponding mRNA. However, in mRNA vaccines, translation of the mRNA occurs rapidly after transfection producing the target protein antigen, which initiates an immune response. For example, mRNA vaccines for COVID-19 use the spike protein-coding mRNA sequence of severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2). Vaccines incorporating mRNA have intrinsically higher safety than those based on DNA, given that mRNA has a short biological half-life and translation takes place in the cytoplasm not the nucleus, reducing the possibility of its genetic information being integrated into the host genome.

3.9 CONTROLLING THE SPREAD OF PATHOGENS AND PARASITES

Controlling the spread of an infectious agent is a complex process, requiring several interventions, some of which will depend on the microorganism involved. Epidemiological investigations (*Chapter 1*) can help to pinpoint the source of the infection. The appropriate disposal of soiled tissues and disposable nappies and avoiding sharing personal items like toothbrushes and shaving razors is also of obvious importance. Strict hygiene precautions, for example thorough cleaning and disinfection of contaminated materials and hand washing, can prevent the spread of pathogens between patients. The boiling of drinking water may be necessary, depending on the source of the outbreak. Routine microbiological investigations may be carried out to ensure that the microbe is under control. For the patient, antibiotics or antiviral drugs (*Section 3.11*) may be given, while patient contacts may be offered vaccination and/or drug treatment. Examples of this multifactorial approach can be seen following outbreaks of *N. meningitidis* in schools and colleges, where all contacts with infected patients are offered antibiotics and vaccines, depending on the serotype. More recently, similar protocols were imposed by many governments in response to the 2019–2023 COVID-19 pandemic.

3.10 INVESTIGATING INFECTIOUS DISEASES

Infectious diseases can affect any organ or system and can cause a wide variety of symptoms and signs (*Chapter 1*). Two criteria must be satisfied to diagnose an infectious disease:

1. signs and symptoms compatible with the suspected infectious agent must be apparent
2. the pathogen must be recovered from the infected site of the patient or there must be evidence of the presence of the pathogen, for instance virulence factors or toxins

Thus, the clinical history and examination of the patient must have the aim(s) of identifying the site(s) of infection and the causative organism. The history should focus on aspects relevant to infectious disease, such as recent travel history, food and water intake, occupational exposure, sexual activity and any use of intravenous drugs. The clinical examination involves identifying fever, skin rashes and swollen lymph nodes (lymphadenopathy) and investigations of the eyes, ears, mouth and throat. Fever is a typical symptom of infection, but not all patients with fever have an infection and not all infectious diseases present with fever. Investigations of the penis or vagina and the throat and rectum are necessary in diagnosing STDs.

Chapter 2. The WHO has also embarked on a worldwide vaccination programme to eradicate polio, but this has not yet been achieved. Of the three strains of wild poliovirus (types 1, 2 and 3), types 2 and 3 were eradicated in 1999 and 2020, respectively. Wild poliovirus type 1 was still endemic in Afghanistan and Pakistan in 2022.

Unfortunately, vaccines against eukaryotic pathogens and parasites are far less developed than those against viruses and bacteria. For example, parasitic diseases caused by protozoa, such as malaria and leishmaniasis, are major public health challenges in tropical and subtropical regions. Given that patients gain natural immunisation to these diseases following recovery, vaccination is a feasible strategy to prevent them. The development of malaria vaccines is discussed in *Box 2.4* but, to date, no vaccines for leishmaniasis have been approved for human use.

Since Jenner's time, vaccination has been widely applied to protect people from many serious infections. Indeed, immunisation is the foundation of primary healthcare systems; safe and effective vaccines are available that prevent or mitigate more than 20 life-threatening diseases such as diphtheria, influenza and meningitis.

The history and physical examination is often supported by clinical tests to assess health and identify the organs affected. These tests include analyses of blood, plasma, urine, sputum, cerebrospinal and seminal fluids and, in some cases, even sweat and solid or semisolid material such as stools. Imaging techniques such as X-rays, ultrasound, computerised tomography (CT) and magnetic resonance imaging (MRI) are assuming greater utility in diagnosis (*Chapters 13* and *18*).

Blood tests involve full blood counts, where eosinophilia is a clinically significant finding with reference to parasitic infections and lymphocytosis is usually indicative of viral infections. C-reactive protein concentration and the erythrocyte sedimentation rate (*Chapters 4* and *13*) are nonspecific tests used to monitor the course of infectious diseases. Other tests involve assessing renal and liver functions (*Chapters 8* and *11*), which may be disrupted by an infection.

Microbiological investigations to identify the infectious agents are especially helpful. They use a variety of techniques, including culture, serological, biochemical and molecular biological tests.

Culture

All human viruses and some microorganisms, such as chlamydiae, are intracellular pathogens and their growth requires the inoculation of cultured eukaryotic cells. When viruses infect and replicate within cultured cells, pathological changes are produced that are characteristic of particular viruses and can be used for identification purposes. Moreover, electron microscopy of supernatants from these cultures, or even on patient biopsy samples, can show the presence of and identify specific viruses (see, for example, *Figure 2.8* of rotavirus virions, which cause outbreaks of diarrhoea in children).

Suspected microorganisms in patient samples are cultured usually on or in growth media, such as nutrient agar or broth, or in selective media that support the growth of specific organisms until growth is detectable. Identification of the organism is based on the shape, texture and colour of colonies and their appearance when stained and observed by microscopy. To some extent, bacteria can be identified by their ability to grow in specific media, such as blood or Sabouraud agars. Indicator media that include some substance that visibly changes because of the metabolic activities of particular microorganisms are also of use in identification. The diagnosis of mycoplasmas, and protozoa and fungal infections can be investigated in similar ways. However, they require a greater use of microscopic and colonial morphology for identification. Helminth parasites can be identified directly from specimens isolated from infected patients and/or indirectly by examining the cysts or eggs of the parasite.

BOX 3.5 MICROBIOLOGICAL TEST STRIPS

Test strips, such as the API types (named from the parent company Appareils et Procédés d'Identification), are widely used to identifying bacteria and yeasts. They were first developed in the 1970s based on recognised biochemical tests but in standardised and miniaturised versions.

A test strip consists of a number of separate compartments each containing the dehydrated reagents necessary for a specific test. The test is typically performed by forming a homogeneous suspension of the microorganism to be identified in 0.85% NaCl solution. Samples of the suspension are then added to each compartment, which rehydrates and activates the reagents. The organisms will produce some observable change in the wells, for example colour changes due to pH differences or enzyme activities or form end products that can be identified. Any one well will, of course, give a positive (+) or negative (–) result. Thus, any one API strip gives a sequence of positive and negative test results that is a numerical identifier (*Figure 3.32*). The organisms can then be identified from a codebook that correlates this sequence with the bacterial species or strain or, more usually, identifications can be made with a computerised database. Organisms can generally be identified accurately and reliably in 4–48 hours, depending upon the strips used and the species of microorganism concerned.

Test strips are available to identify almost all groups of bacteria and many hundreds of different species, including Gram-negative and Gram-positive bacteria, such as species and subspecies of Enterobacteriaceae, bacilli, *Campylobacter*, Corynebacteria, enterococci, *Listeria*, micrococci, staphylococci, streptococci and some anaerobic bacteria and yeasts.

Figure 3.32 A developed API strip showing positive (coloured) and negative (clear) results for each of the test wells in a sequence that is typically indicative of *Salmonella* bacteria.

Serological tests

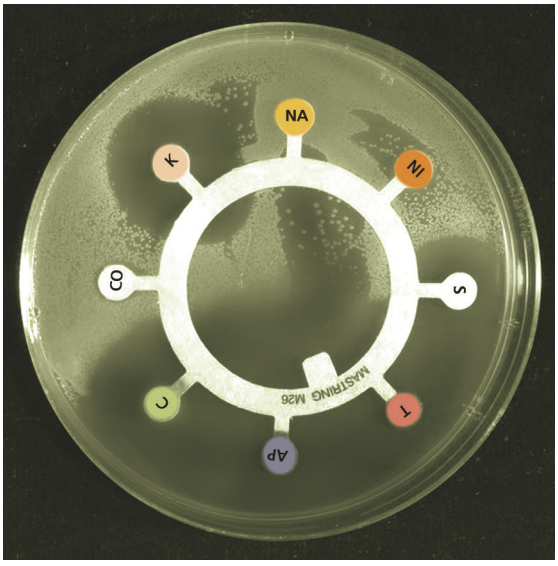
Serological tests involve identifying infectious agents indirectly by measuring antibodies in serum samples from the affected individual. Given that antibodies against a

pathogen increase in number during the early stages of the disease and then decrease during recovery, such tests are particularly useful when it is not possible to isolate the infectious agent. It is preferable to take a blood sample early in the infection (acute serum) and 10–14 days later (convalescent serum). A fourfold or greater increase in antibody titre in the second serum is diagnostic of an acute infection. Antibodies to bacteria can be detected by their ability to agglutinate these microorganisms. Alternatively, a variety of immunotechniques are available, which are outlined in *Chapters 4* and *17*.

Biochemical and molecular biological techniques

A number of biochemical and molecular biological tests are available to help identify microorganisms. The simple Gram stain can immediately eliminate many possible bacteria. The abilities of some bacteria specifically to ferment some carbohydrates, use different substrates, express restricted enzyme activities or form specific products in enzyme-catalysed reactions can all aid in identification.

The susceptibility of bacteria to different antibiotics (*Figure 3.33*) can also be useful and has the added utility of indicating possible therapy.



AP	Ampicillin	NA	Nalidixic acid
C	Chloramphenicol	NI	Nitrofurantoin
CO	Colistin sulphate	S	Streptomycin
K	Kanamycin	T	Tetracycline

Figure 3.33 A multidisc containing a different antibiotic in each lobe as indicated, which can assess the antibiotic susceptibility or resistance of a bacterium. The bacterium in this case is a strain of *Escherichia coli* that cannot grow in the area surrounding an antibiotic to which it is susceptible.

Molecular biological methods are advancing continually. Restriction fragment analyses are described in *Chapters 15* and *16*. The entire genomes of numerous viruses, pathogenic bacteria and parasitic protozoa have been sequenced and the number is growing rapidly thanks to second-generation sequencing methods (*Chapter 17*). Specific genes from many pathogenic viruses and organisms have been cloned and sequenced. That for the 16S ribosomal RNA has received considerable attention and, indeed, differences in the sequences of this gene have allowed evolutionary relationships between different groups of bacteria to be deduced. Nucleic acid probes can be designed to detect characteristic sequences of pathogens and identify the organism in body fluids or tissues particularly using the **polymerase chain reaction** (PCR).

The polymerase chain reaction

The PCR was devised by Kary Mullis (1944–2019) in 1983, for which he was awarded the Nobel Prize in Chemistry in 1993. The shortness of time between the two dates is recognition of the importance of the PCR. Indeed, it is hard to exaggerate the revolutionary impact of the PCR on molecular biology. The PCR is an elegantly simple *in vitro* method for exponentially increasing the number of relatively short strands of DNA of specific sequences, normally 500–5000 bp long, although longer lengths of up to 40 000 bp can also be amplified. These may be whole genes but are more usually only fragments of

DNA. These consist of two polydeoxynucleotide strands arranged in the now familiar double-helical structure. The strands run in opposite directions: one in the 3′→5′ direction, the other with a 5′→3′ orientation. The bases of the nucleotides of each strand associate together to form complementary pairs, with an adenine (A) of one strand paired with a thymine (T) of the other, and guanine (G) pairing with cytosine (C). The base pairs, and therefore the polynucleotide strands, are held together by hydrogen bonds. It is the sequence of bases in the DNA strand(s) that is unique and specific to a particular gene.

The PCR is used to replicate the original DNA sample (template DNA) using a DNA-dependent DNA polymerase, usually abbreviated to DNA polymerase. This enzyme copies the template DNA to produce new strands with complementary sequence. Some DNA polymerases can proofread, that is, correct any mistakes in the newly formed strand, to ensure the fidelity of its sequence. Crucially, DNA polymerases cannot begin the new strand *de novo*, but can only extend an existing piece of DNA. Thus, two artificial oligodeoxynucleotides consisting of fewer than 50 deoxynucleotides are required as primer molecules to initiate the copying process. The primers are complementary to regions that flank the section of the template DNA of interest. Hence, the primer determines the beginning of the region to be amplified. Primers are usually made to order by commercial suppliers that must, of course, be supplied with the required sequence.

The PCR consists of a series of cycles, each of which consists of three identical steps (*Figure 3.34*). First, the double-stranded DNA is heated to 90–96°C to break the hydrogen bonds and allow the two strands to separate. This denaturation is called melting. Prior to the first cycle, the DNA is often heated for an extended time, called a 'hot start', of 5–10 minutes to ensure that the template DNA and the primers are fully melted. In subsequent cycles, 0.5–3 minutes at 94–96°C is normally sufficient for melting. Placing a heated lid on top of the reaction tubes or a layer of oil on the surfaces of the reaction mixtures prevents evaporation. The second step is primer annealing at temperatures of 50–65°C for 30–60 seconds, during which the primers bind to their complementary bases on the now single-stranded DNA templates. The primers must be present in excess of

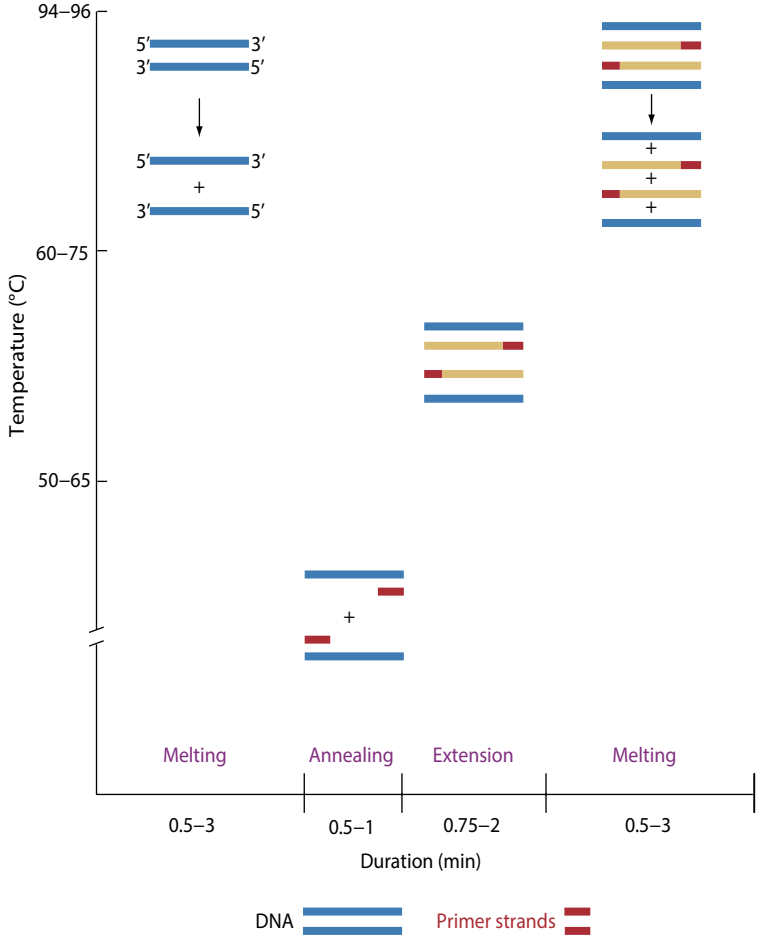


Figure 3.34 Outline of the steps associated with a cycle of the polymerase chain reaction (PCR). ⏎

the target DNA otherwise its strands will simply rejoin. The design of the length of the primers requires careful consideration. Primer melting temperature, which must not be confused with that of the template DNA itself, increases with the length of the primer. The optimum length for a primer is generally 20–40 nucleotides, giving melting temperatures of 60–75°C, although this depends upon their G/C content. If the primers are too short, they will anneal at random positions on the relatively long template and result in nonspecific amplification. However, too long a primer and the melting temperature would be so high, that is, above 80°C, that the DNA polymerase would have reduced activity. The third step of the PCR is the synthesis of DNA in reactions catalysed by the polymerase that extend the primers using the complementary strands as templates. This usually requires 45–120 seconds at 72°C. The extended portion of DNA is a complement of the template strand. Given the high temperatures of the reactions, DNA polymerases from thermophilic organisms are preferred. The *Taq* polymerase from *Thermus aquaticus* is widely used, although it has the disadvantage of lacking proofreading capabilities and therefore of introducing errors (mutations) of 1 in 400–500 nucleotides in the new DNA. Polymerases such as *Pwo* or *Pfu*, obtained from Archea, have proofreading mechanisms that significantly reduce mutations and are used in 'long range' PCR of up to about 30 000 bp.

The result of the first cycle is two helices that are usually overextensions of the target sequences. Each is composed of one of the original strands plus its newly assembled complementary strand and its associated primer, for each of the original template DNA molecules. The cycle is usually repeated 20–30 times in identical conditions with a doubling of the amount of DNA present for every cycle (*Figure 3.35*). Hence, after 30 cycles, the amount of DNA in the original sample has increased over 2^{30} (10^9-fold)! A PCR experiment normally terminates with a 10-minute incubation at 72°C to ensure that all of the amplified DNA molecules are fully extended by the polymerase. Note that extremely stringent conditions are necessary to prevent any unwanted DNA contaminating PCR experiments, since this would also be amplified.

The great advantages of the PCR are the increase in sensitivity and its automation. Samples of DNA from a single cell or from samples many years old, even those embedded in microscope slides, can be amplified and then analysed. The PCR reaction is automated in a thermocycler, which automatically heats and cools the reaction tubes or plates to the appropriate temperatures, for the desired times and the required numbers of cycles. A variation, the reverse transcription PCR (RT-PCR), uses reverse transcriptase to produce a DNA template from an RNA source that can then be amplified by the PCR.

The products of the PCR are identified by determining their base sequences and/or their sizes using agarose or polyacrylamide gel electrophoresis against a known sample. The size of the product can be estimated by comparison with the electrophoretic mobility of DNA fragments of known size (see *Figure 3.36*).

The PCR is used in four major areas of the biomedical sciences:

1. the identification of infectious disease organisms for diagnostic purposes
2. detecting mutations that cause hereditary diseases (*Chapters 15* and *16*)
3. detecting acquired mutation associated with some cancers (*Chapter 17*)
4. tissue typing prior to transplant surgery (*Chapter 6*)

The PCR is one of the most valuable tools for diagnosing infectious diseases. A variety of clinical specimens, such as serum, plasma, urine, semen, cerebrospinal fluid, saliva, and nasal, oral, cervical and ulcer swabs, can be analysed to identify pathogens and parasites. It is especially useful in diagnosing diseases caused by organisms that are difficult or impossible to culture. Its ability to amplify minute amounts of DNA means that PCR-based tests can identify sources of infection accurately, reliably, rapidly and cheaply. Polymerase chain reaction tests can even distinguish the specific strains of papillomaviruses that predispose individuals to cervical cancer. Diagnostic tests are also available for the viruses involved in HIV–AIDS, viral hepatitis and viral meningitis. The PCR is the most sensitive and specific test for *H. pylori*, the main cause of stomach ulcers (*Chapter 11*). Bacterial infections in middle-ear fluid from children suffering otitis media are detectable by the PCR, indicating an active infection, even when standard culture methods fail. The Lyme disease bacterium, *B. burgdorferi* (see *Figure 2.44*), is often difficult to diagnose accurately based on general symptoms, but the PCR can amplify its DNA in body fluids.

Since 1985, many PCR amplification-based techniques have been developed and used clinically. These include multiplex PCR and real-time PCR. Multiplex PCR is useful for diagnosing multipathogenic infections. It uses different pairs of primers to investigate several pathogens, for example the sexually transmitted pathogens *Chlamydia trachomatis*, *Neisseria gonorrhoeae*, *Mycoplasma genitalium* and *Trichomonas vaginalis*, simultaneously in a single specimen. However, it has some limitations, for example

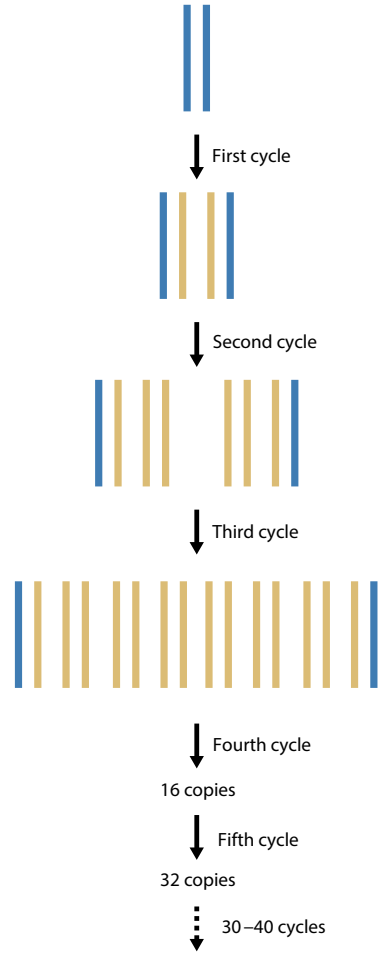


Figure 3.35 The amplification of DNA molecules during cycles of the polymerase chain reaction (PCR).

Margin Note 3.4
The PCR and Hubert H. Humphrey

The extreme sensitivity of the PCR means that it can even be used to diagnose diseases using old samples that are therefore largely degraded. The former USA Vice President and presidential candidate Hubert H. Humphrey (1911–1978) underwent tests for bladder cancer in 1967. However, these gave negative results and, untreated, he died of the disease. A urine sample taken in 1967 and a tissue sample from his cancer-ridden bladder obtained in 1976 were analysed retrospectively with PCR amplification in 1994. The DNA of both samples showed identical mutations in the *p53* gene (*Chapter 17*). This is a well-established tumour-suppressing gene. Had it been possible to diagnose the cancer in 1967, then Humphrey could have received the benefits of the then current treatments and his life may have been extended. This is, of course, only one of many examples of how advances in biomedical sciences have greatly improved clinical practice.

nonspecific products generated by primer–primer interactions can interfere with the amplification of targets.

Currently, real-time RT-PCR is considered the gold-standard test for diagnosing viral diseases, such as Ebola and COVID-19, and detecting bacterial pathogens. In conventional PCR, of course, the amount of accumulated product is assessed at the end of the set number of cycles. However, in real-time PCR, the amounts of product are continuously monitored over the course of the analysis by including a reporter molecule in the initial reaction mixture. Reporter molecules are normally dyes that fluoresce when they intercalate in the nucleic acid product or fluorescent dyes attached to oligonucleotides that are complementary to part of the product's base sequence. Thus, increases in the amount of fluorescence over the course of the amplification, that is, in real time, indicating the amount of product being formed. The cycle threshold (Ct) is defined as the number of cycles required for the fluorescent signal to increase above a set value, that is, exceed the background level (*Figure 3.36*). Cycle threshold values are inversely proportional to the amount of target nucleic acid in the sample (*Figure 3.36*). The speed and sensitivity of real-time PCR means it has numerous clinical uses, such as detecting mutant genes, determining the number of copies of a gene in transformed cells and in analysing the phenotypes of cancer cells (*Chapter 17*).

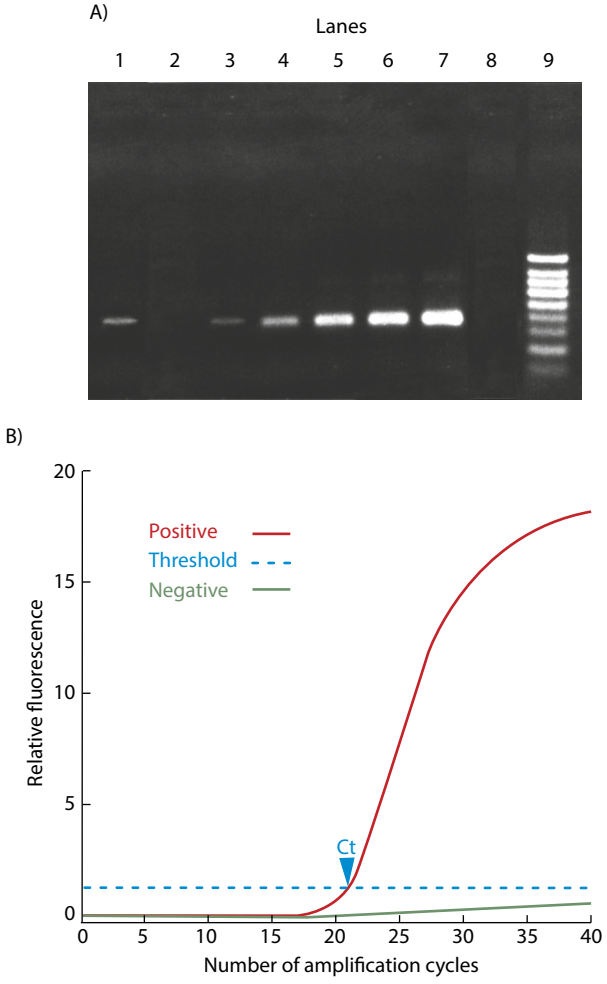


Figure 3.36 (A) A polyacrylamide gel electrophoresis of the products of a polymerase chain reaction (PCR) experiment. Lane 1 shows the DNA sample of 390 bp (positive control), which has been amplified by 10, 20, 25, 30, 35 and 40 cycles of the PCR in lanes 2 to 7. Lane 8 shows the results of a negative control in which the DNA sample was replaced by water and amplification did not occur. Lane 9 shows the separation of DNA molecules of known sizes between 100 and 1000 bp as a calibration marker. (B) The use of real-time reverse transcription PCR to test for coronavirus disease 2019 (COVID-19) in patient samples. One shows a patient is positive, that is, infected (cycle threshold (Ct)=21.4) with severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2), but the other is negative (Ct>40). ⏎

3.11 TREATMENT OF INFECTIOUS DISEASES

The treatment of infectious disease is almost entirely pharmacological, although nursing care is obviously necessary in many cases and surgery may sometimes be used. The development and extensive uses of antibiotics and other drugs over the last 80-or-so years in particular has had an enormous impact in reducing the number and effects of infectious diseases. Unfortunately, natural selection means that all types of pathogens eventually become resistant to drug regimens.

Viruses and bacteria are the most important causes of infectious diseases in developed countries, while fungi, protozoa and helminths are of increased importance in the tropical climates of developing countries.

Antiviral drugs

Viruses are acellular and can only replicate by utilising the metabolic processes of their host cell. It is difficult to target viruses inside host cells without damaging both infected and normal host cells. Indeed, until the development of molecular biology in recent decades there were relatively few effective antiviral agents.

To be fully effective, antiviral drugs should ideally inhibit viral replication but not affect the activities of the host cell. However, many of the current antiviral drugs do not effect a cure but, fortunately, are effective in controlling the disease. Unfortunately, antiviral drugs all interfere to some extent with the reproduction of host cells and so produce adverse effects. Therefore, they must be administered for sufficient time to allow natural immune mechanisms to eradicate all the virions present. Thus, antiviral therapy may well fail in severely immunocompromised patients.

Most antiviral agents act by disrupting one of the steps in the replication cycle of the target virus (see *Figure 2.9* and *Box 3.3*). They may prevent viral internalisation into the host cell. If the virion enters its host cell, other antiviral agents can interfere with the release of the viral nucleic acid from the capsid. Some antiviral agents prevent the synthesis of the viral nucleic acid or its proteins. Even if the virus is successfully synthesised, agents that interfere with its release from the host cell can prevent its dissemination. Lastly, some antivirals can help promote more effective immune response against viral infection. *Table 3.4* lists some of the antiviral drugs used to treat HIV–AIDS, while *Table 3.5* gives examples of ones used for other viral infections.

Margin Note 3.5 ⏎ Combination therapy

Combination therapy is the treatment of infections with two or more drugs, usually to increase the clinical efficacy of the treatment or to minimise the development of resistant strains of the infective organism. With mixed infections involving two or more known pathogens, it is desirable to target each microorganism with one or more different treatments. When the infection is of unknown cause, multiple therapies are advisable to fight the most likely pathogens. Combination therapy may be used even if only a single infective pathogen is present, as using combinations of drugs can prevent or delay the development of resistance to the drugs being used (see *Box 3.2*). Examples of combination therapies include HAART regimens for treating HIV–AIDS (see *Box 3.3*); a quinupristin–dalfopristin cocktail licensed for treating severe Gram-positive infections; the use of isoniazid, rifampicin and pyrazinamide to treat TB; the use of 5-flucytosine combined with amphotericin B to combat some fungal infections; and the combination of albendazole with diethylcarbamazine or ivermectin to treat some helminth infestations (*Section 3.11*).

Antiviral drug	Disease/condition/virus treated
Acyclovir	HHV-1, -2 and -3 (see *Box 3.1*)
Amantadine	Influenza
Brivudine	HHV-1 and -3
Cidofovir	CMV
Entecavir	Hepatitis B virus
Famciclovir	HHV-1, -2 and -3
Foscarnet,	CMV, HHV-1 and -2
Ganciclovir	CMV
Interferon α	Hepatitis B and C viruses
Oseltamivir	Influenza
Penciclovir	HHV-1, -2, -3 and -4
Ribavirin	Adenoviruses, virus poxviruses, respiratory syncytial virus
Rimantadine	Influenza
Vidarabine	Herpesviruses, poxviruses, rhabdoviruses
Zanamivir	Influenza

CMV, cytomegalovirus; HHV, human herpes virus.

Table 3.5 Examples of antiviral drugs. See also *Table 3.4*. ⏎

Viruses can become resistant to specific antiviral agents. Hence, a therapy that combines several agents (*Margin Note 3.5*), each acting at a different stage in the replication cycle, is likely to be a more effective treatment. An example of this is the highly active antiretroviral therapy (HAART) regimens for treating HIV described in *Box 3.3*.

Antibacterial drugs

Drugs to treat bacterial diseases are generally **antibiotics**. They are the most significant in clinical practice because bacteria are responsible for the greater proportion of infectious diseases. Most antibiotics are derived from the products of the metabolism of some bacteria or fungi. Their actions usually rely on differences between microbial and patient cells. Ideally, antibiotic drugs should kill the target pathogen or parasite, that is, they should be **bactericidal** in the case of bacteria. **Bacteriostatic** agents prevent their replication. The latter must be administered for sufficient time to allow the immune defences of the body to eliminate the pathogen.

Not all antibiotics are effective against all pathogenic bacteria. Some are only effective against Gram-negative bacteria, whereas others are effective against Gram-positive only (see later). Some antibiotics can be used to treat a range of both Gram-positive and -negative organisms and are referred to as **broad-spectrum antibiotics**. Antibacterial drugs generally act in one of four major ways: they may inhibit the biosynthesis of the bacterial nucleic acid, proteins or the cell wall or they may act in a variety of rather miscellaneous ways and, in some cases, by unknown mechanisms.

Margin Note 3.6
Antibiotics and toxicity

Despite their specificity, some antibiotics can be toxic to humans, such as the stimulation of allergic responses by penicillins and sulphonamides, and ear and kidney damage by aminoglycosides (*Box 3.6*). In addition, chloramphenicol can show liver and bone marrow toxicities causing serious haematological diseases, particularly aplastic anaemia. This can cause the death of rare, susceptible individuals.

Inhibitors of DNA replication and transcription

Many drugs inhibit both replication and transcription at several points. These drugs fall into several general families, for example the:

- nitroimidazoles
- sulphonamides or sulpha drugs
- quinolones and fluoroquinolones
- rifamycins

Nitroimidazoles, such as metronidazole, are nonspecific in action and bind directly to DNA, denaturing its helical structure such that it is no longer a substrate for DNA-binding proteins. Nitroimidazoles display broad-spectrum activity against anaerobic Gram-positive and -negative bacteria, mycobacteria and parasites. Although they have been used since the 1950s, there is now renewed interest in the clinical use of new derivatives, particularly to treat TB and parasitic infections. Clofazimine is used to treat leprosy and multidrug-resistant TB. It is not a nitroimidazole but it also binds to DNA, preventing replication and transcription. However, its mechanism(s) of action is (are) complex and poorly understood. Studies have indicated that it induces apoptosis with fragmentation of DNA (see *Box 17.3*). It also appears to interfere with oxidative phosphorylation and stimulates bacterial phospholipase A2, increasing the release of lysophospholipids, an action that is toxic to mycobacteria.

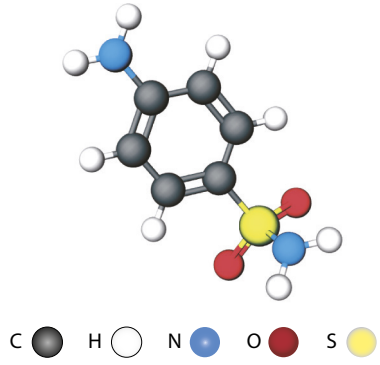


Figure 3.37 The sulpha drug, sulphanilamide.

Sulphonamides or sulpha drugs are an extensive group of antibiotics, which, since their introduction in the 1930s, have seen continual clinical use (*Figure 3.37*). They are bacteriostatic, inhibiting the formation of folate: an essential precursor in the biosynthesis of nucleic acids. Sulphonamides displayed clinical activity against Gram-positive and -negative bacteria, but the development of bacterial resistance has led to decreased use. They are, however, still used to treat UTIs and protozoal diseases, in both instances as an ingredient of combination therapies.

Quinolones are widely used broad-spectrum antibacterial drugs. The initial first generation, nalidixic acid has been superseded by second-, third- and fourth-generation fluoroquinolones, for example ciprofloxacin, pefloxacin and delafloxacin, respectively. More than a dozen are now in clinical use. They are used to treat Gram-positive and -negative bacterial infections such as gastroenteritis, gonorrhoea, skin and soft-tissue infections, TB and infections of the urinary tract.

Both quinolones and fluoroquinolones bind to DNA gyrase and DNA topoisomerase IV; both are bacteria type II topoisomerases, that is, they catalyse the hydrolysis of both strands of DNA. They have related and overlapping roles in bacterial DNA replication. Gyrase catalyses the ATP-dependent negative supercoiling of the circular bacterial DNA molecule by cleaving both strands of the DNA and then twisting the two ends around each other. Topoisomerase IV separates the topologically linked daughter chromosomes during the late stages of DNA replication. The binding of quinolones/fluoroquinolones stabilises the gyrase–DNA and topoisomerase IV–DNA complexes forming drug–enzyme–DNA complexes in which both strands of the DNA are severed (*Figure 3.38*). This initially results in bacteriostasis but the release of DNA with double-stranded breaks

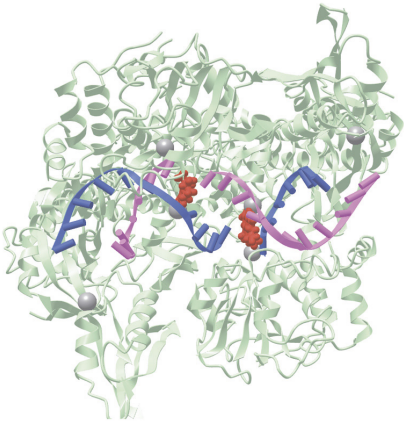


Figure 3.38 Molecular model of the topoisomerase IV of *Streptococcus pneumoniae* with a fluoroquinolone (red) bound at its active site. Note the cleaved DNA (shown in blue and magenta). PDB file 3RAD.

from the gyrase and/or topoisomerase IV complexes results in chromosome fragmentation and pathogen death.

Fluoroquinolones have been used clinically for over 30 years, and both moxifloxacin and levofloxacin (a second-generation fluoroquinolone) are still used to treat multidrug-resistant bacteria. Ciprofloxacin continues to be widely used for a variety of diseases caused by Gram-negative and Gram-positive bacteria, including respiratory and urinary infections and several enteric diseases. However, the last 20 years have seen increases in strains of bacterial resistance to current fluoroquinolones, and alternative compounds are urgently needed. Also, some, for example grepafloxacin and sparfloxacin, have had to be withdrawn because of serious adverse side effects. Thus, the development of novel antibiotics to treat clinically resistant microbial strains is essential; for example, the fluoroquinolone delafloxacin is one that has relatively recently entered clinical use.

Rifamycins were discovered in 1957. They act by inhibiting bacterial DNA-dependent RNA polymerases, which suppresses transcription. Rifamycins are bactericidal against numerous Gram-positive and -negative bacteria. Those in contemporary use include rifampin, rifabutin and rifapentine. Indeed, although rifampin was introduced over 40 years ago, it is still essential in combination therapies (see *Margin Note 3.5*) to combat drug-resistant bacteria when treating TB (*Figure 3.39*) and the opportunistic *Mycobacterium avium-complex* that invades the pulmonary system of susceptible individuals. *Mycobacterium abscesses* is a distant relative of *M. tuberculosis* and *M. leprae* (*Chapter 2*) and an opportunistic pathogen. It causes difficult-to-treat lung infections in tubercular patients; however, rifabutin has been shown to be active against this pathogen.

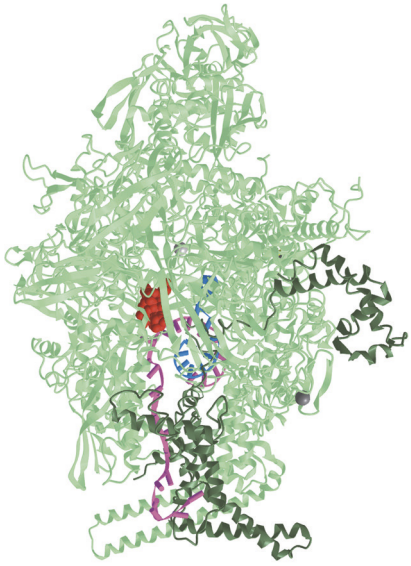


Figure 3.39 Molecular model of the rifamycin antibiotic rifampicin (red) bound to the transcription initiation complex of *Mycobacterium tuberculosis*. The DNA-dependent RNA polymerase and its σ factor are shown in green and dark green, respectively, while DNA is in blue and magenta, and Zn^{2+} and Mg^{2+} are shown as dark- and light-grey spheres, respectively. PDB file 5UHB.

Inhibitors of mRNA translation

The translation of messenger RNA molecules by ribosomes forms polypeptides. Translation is broadly similar in both bacterial and mammalian cells, although there

BOX 3.6 OTOTOXICITY AND MÉNIÈRE DISEASE

Approximately 5–10% of patients treated with the aminoglycosides streptomycin and gentamicin (*Figure 3.40*) experience side effects, involving **ototoxicity**, that is, damage to the inner ear, although the extent varies with the drug and dose. In most cases, damage is minor and reversible once medication ceases. Ototoxicity is obviously undesirable; however, the ear damage associated with aminoglycosides can help some patients who suffer from Ménière disease, named after the French physician Prosper Ménière (1799–1862). Its incidence is thought to be 0.5–7.5 per 1000. It is **idiopathic** and affects the inner ear, being characterised by recurring episodes of hearing loss, tinnitus, rotational vertigo (a form of dizziness), nausea and a sense of pressure in the middle ear. In some patients, a chemical labyrinthectomy involving the injection of gentamicin through the tympanic membrane into the middle ear, from where it diffuses into the inner ear and destroys some or all of the balance cells, may be performed to control the vertigo associated with the affected ear. This procedure causes less damage to the hearing mechanism than some traditional treatments. Ménière patients with symptoms in both ears can be given intramuscular injections of streptomycin.

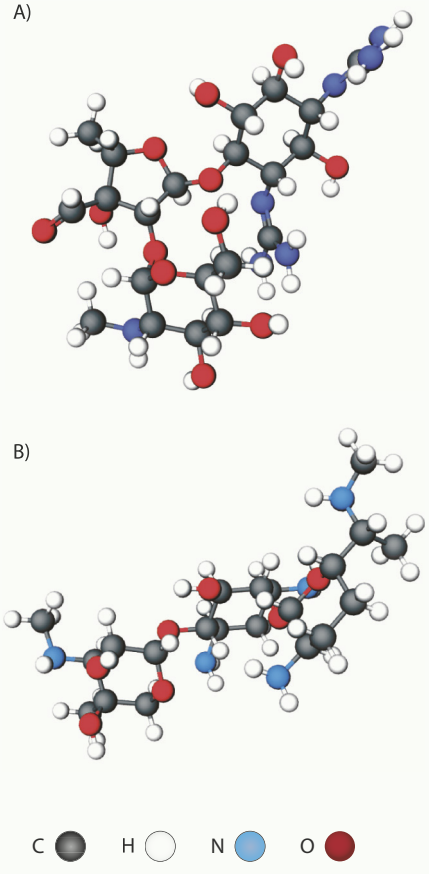


Figure 3.40 Computer drawn models of (A) streptomycin and (B) gentamicin.

are significant differences between the two types of cells. For instance, bacterial ribosomes consist of 30S and 50S subunits, whereas eukaryotic ones have larger 40S and 60S subunits. Several of the protein translation factors necessary for translation also differ. These differences are exploited by antibacterial drugs.

The main antibacterial antibiotics that interfere with mRNA translation are:

- aminoglycosides
- lincosamides
- macrolides
- streptogramin: quinupristin–dalfopristin
- tetracyclines
- miscellaneous/historical

Aminoglycosides, such as streptomycin and gentamicin, are broad-spectrum bactericidal agents. They bind irreversibly to specific proteins of the 30S subunit and inhibit translation by interfering with its initiation and preventing proofreading. This is an essential checking step performed by the ribosome; hence, incorrect amino acid residues are inserted into the polypeptide, leading to the production of nonfunctional or toxic peptides. Aminoglycosides were largely replaced in the 1980s by cephalosporins (see later). However, the increased incidence of multidrug-resistant bacteria means there is renewed interest in using these drugs in combination therapies with other types of antibacterial classes.

Lincosamides are named after Lincoln, Nebraska, where they were first isolated from *Streptomyces lincolnensis*. The major lincosamides are clindamycin and lincomycin, both of which are bacteriostatic in sensitive bacteria. However, at high concentrations, they can be bactericidal. Clindamycin has greater antibacterial activity, particularly against anaerobic bacteria but also some protozoa. It is therefore used to treat some protozoal diseases, for example malaria, usually combined with another antimicrobial compound, such as primaquine.

Macrolides are a large family of broad-spectrum antibiotics that are bacteriostatic or bactericidal. They inhibit the translation of mRNA by binding to the 50S ribosomal subunit preventing elongation and inhibiting ribosomes from forming functional polyribosomes. The first macrolide used was erythromycin, followed by azithromycin and clarithromycin (*Figure 3.41*). More recently, the evolution of resistance to these antibiotics led to the development of a third generation of macrolides, which has increased activity against many macrolide-resistant bacteria. Unfortunately, this was accompanied by serious side effects, although newer macrolides have been synthesised with reduced adverse effects.

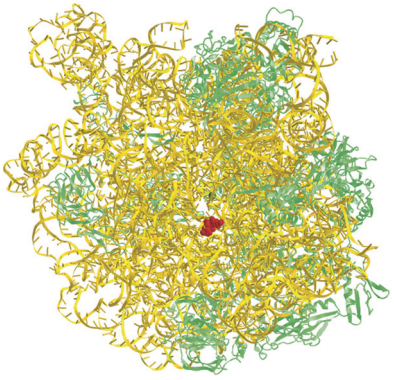


Figure 3.41 Molecular model showing the binding of the macrolide antibiotic clarithromycin (red) to 50S ribosomal subunit from *Mycobacterium tuberculosis*, the causative pathogen of tuberculosis. Ribosomal RNAs and proteins are indicated in yellow and light green, respectively. PDB file 7F0D. ⏎

Tetracyclines are a well-known group of drugs that includes the parent tetracycline (*Figure 3.42*), as well as other antibiotics, such as doxycycline, minocycline and oxytetracycline. Tetracyclines specifically inhibit the binding of an aminoacyl-tRNA to the acceptor site on the mRNA–30S ribosome subunit complex and therefore slow translation. Thus, their actions against aerobic and anaerobic bacteria, as well as Gram-positive and -negative bacteria, with exceptions of *Proteus* spp. and *P. aeruginosa*, are bacteriostatic. However, as with nearly all antibiotics, there is a concern over tetracycline-resistant bacterial strains.

Streptogramins type A and B, dalfopristin and quinupristin, respectively, are produced by *Streptomyces pristinaespiralis* and chemically modified to give the drug quinupristin-dalfopristin. Quinupristin or dalfopristin alone shows weak bacteriostatic activities. However, together their actions are synergistic, since they target different sites on the 50S ribosome subunit. Dalfopristin binds directly within the peptidyl transferase centre of the ribosome interfering with the binding of tRNA-aminoacyl molecules and inhibiting elongation of the polypeptide, while quinupristin blocks the exit tunnel of the ribosome interfering with polypeptide elongation. This inhibition of translation leads to the release of incomplete polypeptides.

Quinupristin–dalfopristin was licensed for use in the UK and USA in the late 1990s for treating severe infections with Gram-positive organisms, including **nosocomial** pneumonia and infections related to the use of intravascular catheters. It is particularly useful for treating complicated skin infections with methicillin-resistant *S. aureus* and *S. pyogenes* and life-threatening infections of vancomycin-resistant *Enterococcus faecium*. Quinupristin–dalfopristin has poor activity against *E. faecalis* compared with the clinically less serious *E. faecium*. However, strains of *E. faecium* resistant to quinupristin-dalfopristin are found.

Several other antibiotics that interfere with mRNA translation are also used clinically. Fusidic acid is a narrow-spectrum steroidal antibiotic first isolated in 1960 from the fungus *Fusidium coccineum*. It inhibits translations by binding to elongation factor G (EF-G), a bacterial protein that is necessary for the translocation step that follows

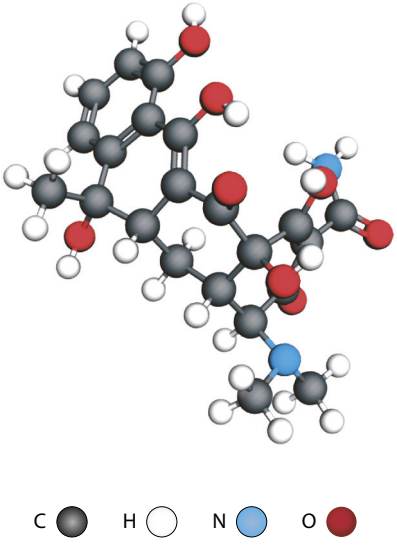


Figure 3.42 Computer drawn model of tetracycline. ⏎

peptide bond formation. It is active mainly against Gram-positive bacteria and is used to treat local and systemic staphylococcal infections, including those caused by methicillin-resistant *S. aureus*. However, resistance strains have emerged and it is now used largely in combination with other antistaphylococcal agents.

Chloramphenicol and related antibiotics were original first-line antibiotics to treat patients with enteric fever. It inhibits the formation of peptide bonds in translation by binding to the peptidyltransferase of the 23S rRNA of the bacterial 50S ribosome subunit (*Figure 3.43*). This action prevents an incoming aminoacyl-tRNA interacting with the A site of the ribosome and results in a bacteriostatic action. Chloramphenicol is no longer routinely used in many areas because of the risk of irreversible bone marrow aplasia (defective development) in up to 1 in 10 000 recipients and because of the evolution of resistant strains of *S. typhi* in the 1980s. However, chloramphenicol is still used in some regions with limited resources because of its broad-spectrum activity against many common causes of serious childhood bacterial infections, including enteric fever.

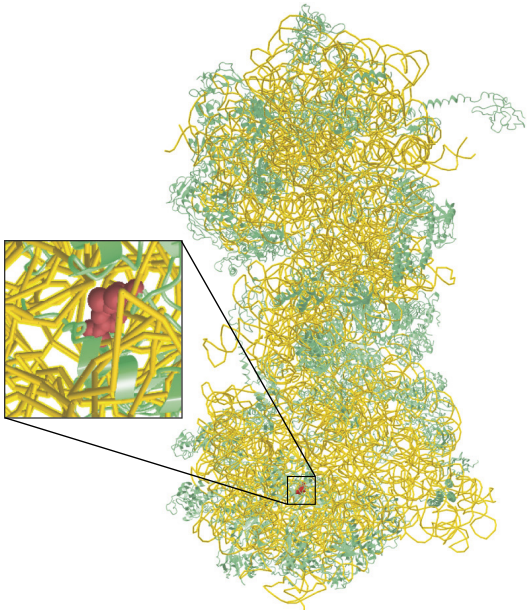


Figure 3.43 Molecular model showing chloramphenicol bound to an *Escherichia coli* ribosome. The ribosomal RNAs and proteins are shown in yellow and light green, respectively. PDB file 4V7T.

Lastly, spectinomycin inhibits bacterial translation of mRNA by binding to the 16S ribosomal RNA and the ribosomal protein termed RpsE, where it interrupts the translocation step of polypeptide biosynthesis. It has been widely used to treat infections, particularly those of penicillin-resistant staphylococci and *N. gonorrhoeae*. However, its effectiveness has declined with the evolution of resistance mechanisms in target organisms. Currently, research efforts to develop chemical analogues, called spectinamides, with effective antibiotic activity are underway.

Inhibitors of bacterial cell wall biosynthesis

Human cells, like all animal cells, lack a cell wall. Hence, metabolic processes concerned with bacterial cell wall biosynthesis are excellent targets for specific antibacterial agents. Bacterial cell walls form protective bags around each microbial cell, which prevents osmotic lysis. They contain layers of peptidoglycan, each of which consists of chains of amino sugars linked together by short peptides. Gram-negative bacterial cell walls have only a single layer of peptidoglycan. In contrast, those of Gram-positive organisms may have as many as 40. Bacterial growth involves cell division and entails the breakdown of the cell wall by bacterial enzymes, followed by the biosynthesis of new peptidoglycan. Antibiotics, such as the β lactams, that inhibit cell wall synthesis are almost all bactericidal, since they stop new peptidoglycan formation and prevent cell wall biosynthesis, while bacterial enzymes continue to degrade the existing cell wall.

β Lactams are the largest, most widely used class of antibacterial antibiotics and contain, perhaps, the best-known antibiotics:

- penicillins
- cephalosporins

They are, of course, only active against growing bacterial cells. All contain the β lactam structural ring (*Figure 3.45*) responsible for their antibacterial effects. β Lactams are irreversible inhibitors of transpeptidase, the enzyme that catalyses the formation of cross links between the sugar residues and peptides in the peptidoglycan layer(s). Bacterial resistance to β lactams has developed in strains previously susceptible to them. This involved the evolution of β lactamase, an enzyme that degrades the drug. Their clinical life has, however, been prolonged by drug formulations that included inhibitors of β lactamase.

Several antibiotics other than β lactams also reduce the efficiency of bacterial cell wall biosynthesis. They inhibit several different enzyme-catalysed steps at a wide variety of sites; hence, they have no common mechanism of action. Examples of such drugs are:

- cycloserine
- fosfomycin
- isoniazid
- vancomycin

Cycloserine is a structural analogue of D-alanine, which has broad-spectrum antibiotic activity. It is used to treat *M. tuberculosis* and *M. bovis*. Cycloserine competitively inhibits L-alanine racemase and D-alanine:D-alanine ligase, whose activities are necessary for the biosynthesis of peptidoglycans.

Fosfomycin was first isolated in 1969 from cultures of *Streptomyces* spp. It is a bactericidal broad-spectrum antibiotic. It is an analogue of phosphoenolpyruvate (*Figure 3.46*), which irreversibly inhibits pyruvyltransferase, an enzyme that catalyses early stages in peptidoglycan biosynthesis. It is used to treat uncomplicated UTIs caused by *E. coli* and *E. faecalis* but is especially useful treating cystitis caused by bacteria that produce extended-spectrum β lactamases that make them resistant to a range of antibiotics. Hence, its use for combating infections with multidrug-resistant Gram-negative bacteria, such as *Enterobacteriaceae*, *P. aeruginosa* and methicillin-resistant *S. aureus*, sometimes in combination with other antibiotics.

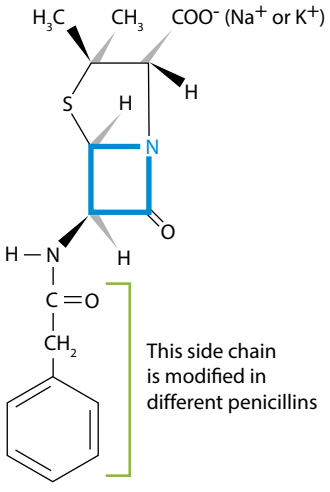


Figure 3.45 Natural penicillin. The β lactam ring is shown in blue.

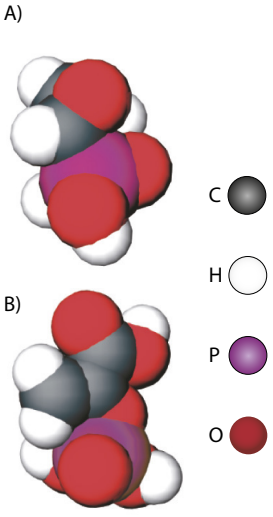


Figure 3.46 Computer drawn models of (A) fosfomycin and (B) phosphoenolpyruvate.

Isoniazid (see *Figure 10.33*) is a synthetic antimicrobial first introduced into therapy in 1952 to treat TB, which the WHO has listed as the largest cause of death by any infectious agent globally. Isoniazid was the first drug demonstrated to have specific bactericidal activity against *M. tuberculosis* and still is in the front line of TB treatment, especially when combined with the antibiotics rifampicin, pyrazinamide and ethambutol, although the treatment must be sustained for an extended period. Isoniazid is a prodrug. *In vivo*, it reacts with NAD^+ (see *Figure 10.13*) to form a complex (*Figure 3.47*) that competitively inhibits NADH-dependent enoyl-ACP reductase. This enzyme catalyses

Margin Note 3.7
Fleming and penicillin

The antibacterial activity of penicillin was discovered by Alexander Fleming (1881–1955) in 1928 at St Mary's Hospital in London. During this period, Fleming went on holiday and left some cultures of staphylococci on agar plates unwashed. When he returned, he noticed that a fungal contaminant growing on one of the plates was inhibiting the growth of the bacteria around it. Fleming realised that the mould was secreting a substance into the agar that prevented the bacteria from growing. He called this unknown substance **penicillin** after the contaminating mould *Penicillium notatum* (*Figure 3.44*). However, Fleming was unable to isolate the substance, and its purification did not occur until many years later following extensive work by Ernst Chain (1906–1979) and Howard Florey (1898–1968) in Oxford during World War II. The significance of Fleming's discovery of the first antibiotic was quickly appreciated and, among many honours, Fleming was elected Fellow of the Royal Society in 1943 and knighted in 1944. In 1945, Fleming, Chain and Florey were jointly awarded the Nobel Prize in Physiology or Medicine for their discovery of penicillin and its curative effects.

Figure 3.44 *Penicillium* mould growing on agar.

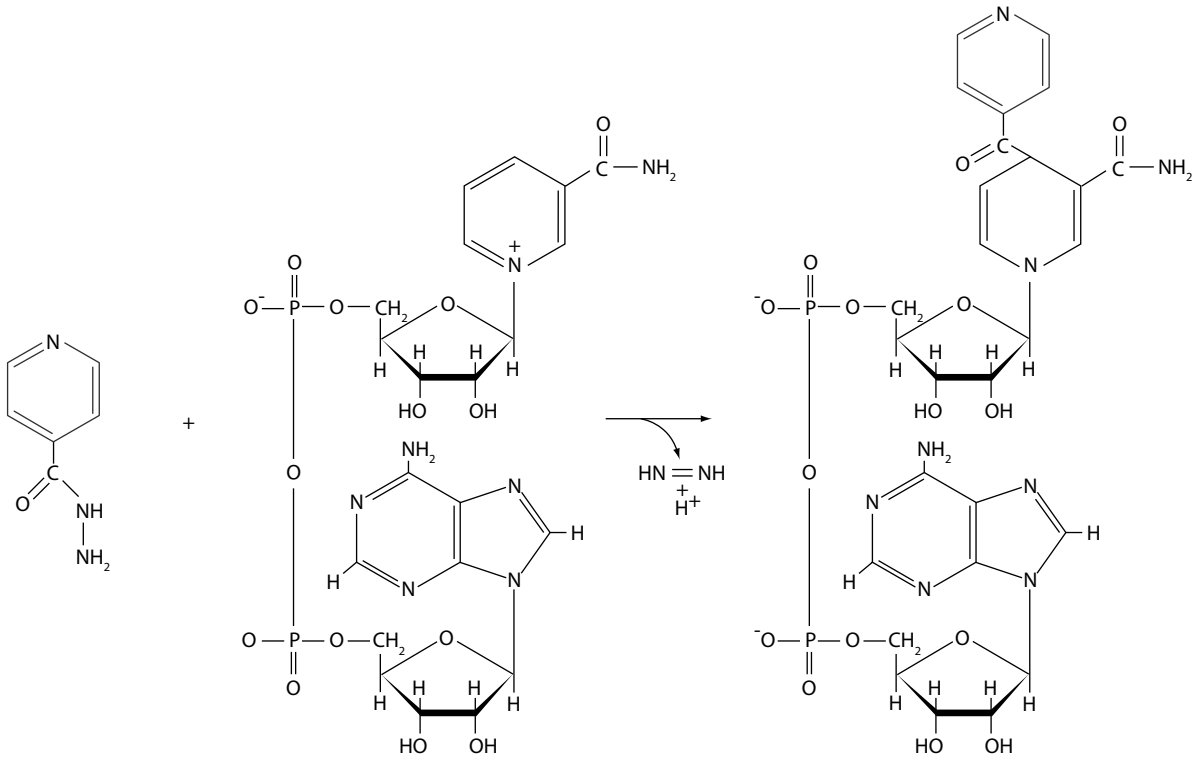


Figure 3.47 The reaction between isoniazid and NAD^+.

an essential step in the biosynthesis of very long chain fatty acids, which are precursors of mycolic acid components of the mycobacterial cell wall. Treatment with isoniazid therefore results in the accumulation of very long chain fatty acids within the bacterium and detrimental morphological changes in its cell wall.

Vancomycin (*Figure 3.48*) inhibits transpeptidase and carboxypeptidase activities involved in cell wall formation. It is a major agent in treating antibiotic-resistant *C. difficile* infections. Unfortunately, resistance to vancomycin in *C. difficile* has been reported.

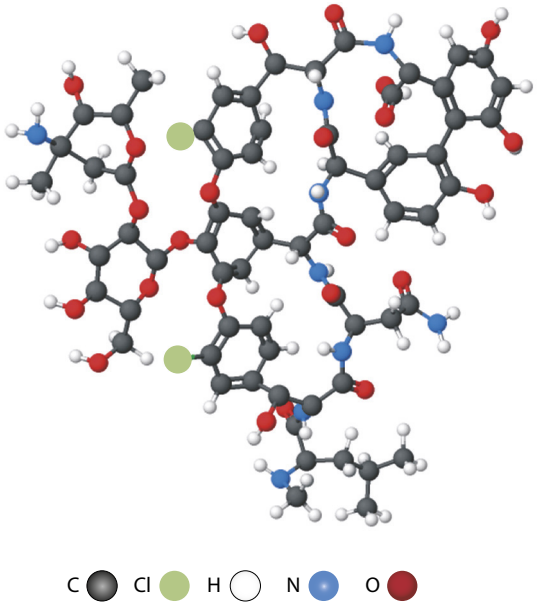


Figure 3.48 Computer drawn model of vancomycin.

Miscellaneous antibiotics

A (rather) miscellaneous collection of antibiotics includes metronidazole, nitrofurantoin, polymyxins and pyrazinamide.

Metronidazole is a bactericidal synthetic derivative of azomycin that was originally found in cultures of *Streptomyces* spp. (see *Figure 3.59*). It was introduced into clinical use in 1959 and still remains the drug of choice, singly or combined with other antibiotics, to treat clinical conditions caused by anaerobic bacteria, such as colon infections with *C. difficile*; infectious microaerophilic organisms, for example the bacterium *H. pylori*; and several protozoan parasites (see below). Metronidazole is also used to treat abscesses of the liver, pelvis, abdomen and brain caused by susceptible anaerobic bacteria of, as well as infections of, the oral cavity, such as gingivitis and periodontitis (see *Section 3.5* and *Box 11.2*).

The cytotoxicity of metronidazole and related compounds is still subject to study. They are prodrugs, and the unionised form of metronidazole is readily taken up by susceptible organisms. Within the cytoplasm, they are metabolically reduced at very low concentrations of oxygen, which explains their cytotoxicity to anaerobic and microaerophilic microorganisms. The metabolic intermediates produced by reduction are highly reactive and target many molecules: they damage DNA, impairing its replication and repair, and form covalently linked complexes with cellular proteins, resulting in death of the pathogen. Common side effects involve gastrointestinal problems. Also, bacterial resistance to metronidazole has occurred, especially in strains of *H. pylori*.

Nitrofurantoin was discovered in the 1940s. It is effective against most Gram-positive and -negative bacteria and used as a short-term first-line therapy for uncomplicated lower UTI or given long term as a prophylaxis against recurrent chronic infections. The mechanism(s) of action(s) of nitrofurantoin is/are relatively poorly understood. It is a prodrug that is reduced by bacterial flavoproteins (*Figure 3.49*) to intermediates that inhibit the biosynthesis of bacterial enzymes involved in DNA replication and transcription and cell wall formation. Nitrofurantoin is one of the most utilised antibiotics to treat bladder infections and UTIs. Toxicity is generally mild, but its use demands vigilant monitoring, given it can damage multiple organs, especially the lungs and the liver. Indeed, it is one of the most common causes of drug-induced liver disease. It has been associated with nitrofurantoin-induced drug reaction/rash with eosinophilia and systemic symptoms (DRESS) syndrome.

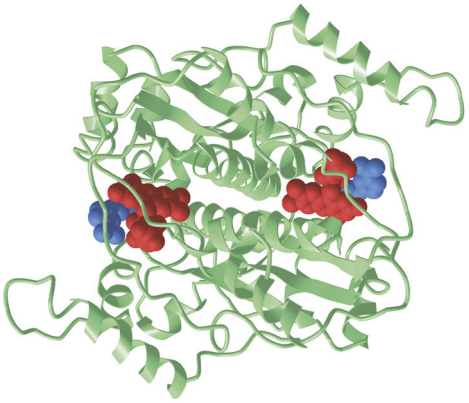


Figure 3.49 Molecular model of an *Escherichia coli* flavoprotein with bound nitrofurantoin molecules (blue) and two FMN molecules (red). PDB file 7NB9. ⏎

Polymyxins are fermentation products of *Bacillus polymyxa* (reclassified in 1993 as *Paenibacillus polymyxa*). Only polymyxins B and E (colistin) are of use clinically. These are amphipathic, detergent-like lipopeptides that are bactericidal to Gram-negative bacteria (*Figure 3.50*). Polymyxins bind to the lipopolysaccharides of the outer membrane of the cell wall in these bacteria, which displaces its stabilising Mg^{2+} and Ca^{2+}. This disruption allows periplasmic proteins to leak away and other molecules to enter the space, including the polymyxins. These can then attack and disrupt the cytoplasmic membrane, destabilising the osmotic balance of the bacterial cell leading to it swelling, lysis and death. Polymyxins are less active against Gram-positive bacteria because they have fewer phospholipids in their outer cell wall and cytoplasm membrane. The membrane attack active mechanism of action for polymyxins is generally accepted but there is evidence for additional mechanisms of action, such as inhibiting bacterial respiration and ATP production.

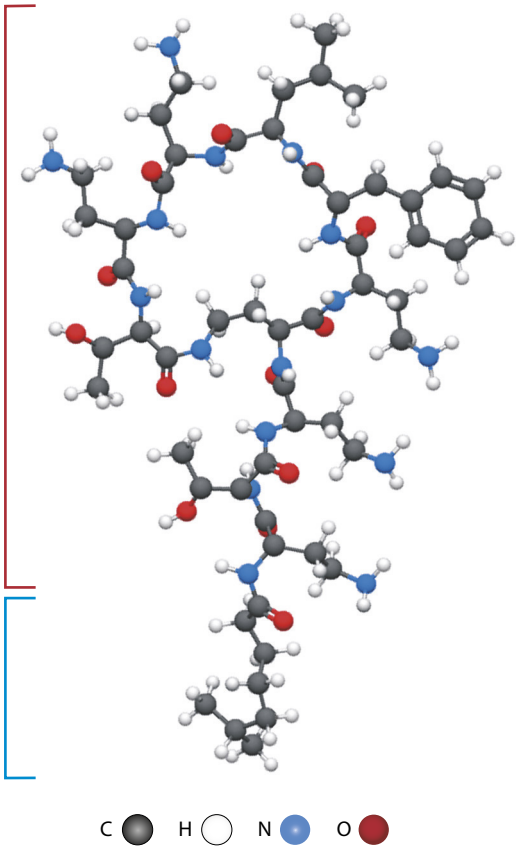


Figure 3.50 Computer drawn model of polymyxin B. Note the hydrophilic and hydrophobic portions of the molecule, indicated by the red and blue lines, respectively, that impart its amphipathic properties.

Polymyxins were commonly used in clinical treatment until the 1980s, but their use declined because of associated nephrotoxicity and neurotoxic effects. However, polymyxin B and colistin have re-entered clinical use to treat multidrug-resistant Gram-negative bacteria, for example *P. aeruginosa*. With more knowledgeable and experienced use, the incidence of toxic side effects is both less common and severe. Unfortunately, multidrug-resistant Gram-negative pathogens have now also become resistant to polymyxins.

Pyrazinamide is a synthetic analogue of nicotinamide (*Figure 3.51*) used in combination therapies with isoniazid or rifampin to treat TB. It is bactericidal against the intracellular pathogen *M. tuberculosis* in acidic environments, for example within macrophage cytoplasm or in inflammatory lesions. Here, the bacteria-specific pyrazinamidase catalyses its conversion to pyrazinic acid, which decreases the pH to values that deplete membrane potentials and inhibit intracellular targets, eventually resulting in death of the pathogen. Pyrazinamide is a first-line anti-TB medication. It is effective

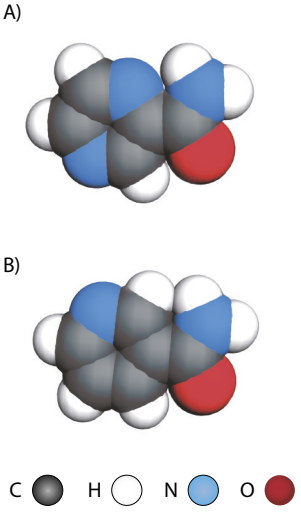


Figure 3.51 Computer drawn models of (A) pyrazinamide and (B) nicotinamide. Note their structural resemblance.

BOX 3.7 BACTERIOPHAGES AND PHAGOTHERAPY

Bacteriophages or, more simply, phages are viruses that can infect and kill bacteria (*Figure 3.52*).

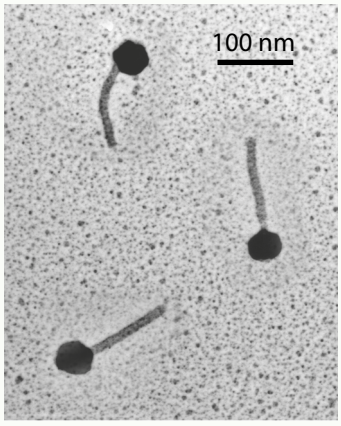


Figure 3.52 Electron micrograph showing three λ virions, a parasitic bacteriophage of *Escherichia coli*. Courtesy of Northwest Regional Virus Laboratory, Booth Hall Hospital, Manchester, UK. ⏎

They were discovered in 1915 by Frederick Twort (1877–1950) and named by Félix d'Hérelle (1873–1949) in 1917. The clinical use of bacteriophages to kill infectious bacteria is called phagotherapy or phage therapy.

Bacteriophages may be separated into two groups:

1. lysogenic or temperate
2. lytic or virulent

Lysogenic or temperate bacteriophages insert their nucleic acid into a susceptible bacterial cell, where it is integrated into the bacterial chromosome. Hence, the viral genome is replicated only when the cell divides. Lytic or virulent bacteriophages also insert their nucleic acid into bacterial cells. However, it is not integrated into the bacterial genome but subverts the biosynthetic apparatus of the cell to produce new virions. These are formed up to a burst number of 50–200 virions, which varies with the type of bacteriophage, triggering the lysis and death of the bacterial cell. The new virions released can then infect new bacterial cells (*Figure 3.53*). It is apparent that only lytic bacteriophages can be of use in phage therapy.

Phagotherapy has been used therapeutically for over 100 years. Indeed, following their discovery, d'Hérelle began almost immediately experimenting with bacteriophages to treat infections, with some success. Interest in phagotherapy declined sharply in the USA and western Europe in the 1940s. The reasons for this decline are debated, but appears mostly to have followed the discovery of antibiotics, starting with penicillin in 1928 and then the development of increasing numbers of safe and effective drugs following World War II.

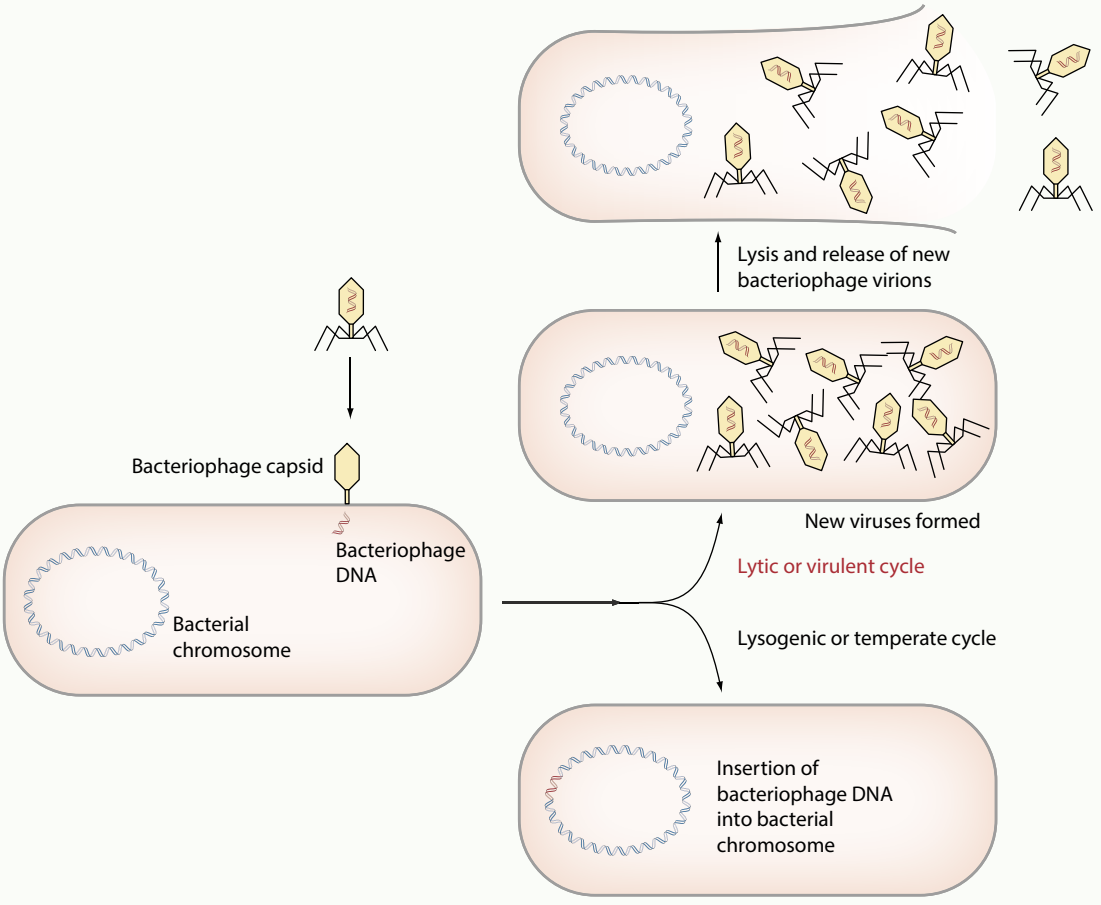


Figure 3.53 Schematic illustrating the two different replication strategies of lysogenic/temperate and lytic/virulent bacteriophages. ⏎

However, phagotherapy was in continual therapeutic use in eastern Europe and especially in Georgia and Russia and was still occasionally practised in Belgium and France.

The evolution of ever-increasing numbers of strains of antibiotic-resistant virulent bacteria has reawakened general interest in the use of bacteriophages as therapeutic agents. Given that one estimate of the number of bacteriophage virions is 1×10^{31}, it is difficult to image a bacterium that cannot be killed by at least one of them. Also, bacteriophages are easy to obtain; any areas prolific in human-associated bacteria, including pathogenic types, are rich sources. This includes soil, sewage and contaminated water. Once isolated, bacteriophages are relatively easy to purify and can be stored frozen.

Phagotherapy offers a number of advantages. Bacteriophages are extremely efficient at killing host bacteria, and their effectiveness can be improved artificially by gene modification (*Chapter 17*). Further, given the narrowness of their host specificities, bacteriophages cannot attack human cells. It also ensures that their effects upon commensal microorganisms are likely to be minimal compared with broad-specificity antibiotics (*Chapter 2*). They can be used alone but also in combination with antibiotics and other drugs, although the bacteriophage(s) need only be applied once because of the increase in their number during treatment. Bacteria, of course, become resistant to these parasites but, because bacteriophages are biological entities, they coevolve and develop into new virulent strains.

All therapeutic approaches have some problematical features; phagotherapy is no exception. Given that a specific bacteriophage will kill only a particular strain of pathogen, it may take weeks to both identify and obtain the required virus. This is short compared with the years required to develop a new antibiotic, but may be a relatively long time in terms of the needs of the patient. Also, in this sense, the very specificity of bacteriophages is a disadvantage, as a bacteriophage will kill only a specific strain of its host bacterium. Consequently, mixtures of different bacteriophages, occasionally referred to as cocktails, are often used to improve the clinical effectiveness of the treatment. A common problem is determining an appropriate dose of the cocktail to use, which, in turn, leads to difficulties in standardising formulations for clinical use.

Clinically, bacteriophages have had three major applications: first, to combat infections involving antibiotic-resistant bacteria; second, to treat infections that fail to respond to antibiotics because of some clinical feature presented by the patient, for example poor circulation or diabetic ulcers; and, third, to eliminate protected bacteria, such as those in a biofilm or anaerobic types encased in a capsule. In these instances, the bacteriophage(s) can be placed directly on the infected site.

Phagotherapy has been extensively and successfully used in topical creams, lotions and ointments containing lytic bacteriophages to combat *Pseudomonas aeruginosa* and *Staphylococcus aureus* infections following wound and burn skin lesions. Skin ulcers affecting feet in diabetic patients have responded well to similar therapies. A major cause of endodontic infections is the bacterium *Enterococcus faecalis*. Endodontic diseases (see *Section 3.5* and *Box 11.2*) are biofilm-mediated infections and are often difficult to treat with conventional antibiotics. Again, in several cases, bacteriophages have successfully eliminated these types of infections. Bacteriophages can also be administered orally. Lung infections of *P. aeruginosa* and the Gram-negative bacterium *Achromobacter xylosoxidans* in cystic fibrosis patients (*Chapter 16*) have been effectively treated by tablet formulations of bacteriophages. Cholera and dysentery cases have also responded well to tabulated bacteriophages.

The use of bacteriophages in clinical therapy appears attractive. They have a long history of use in some countries, and successful clinical cases have been reported in Europe and the USA and received a great amount of publicity. However, only a few, large-scale appropriately designed randomised controlled clinical trials to evaluate phagotherapy have been reported. Hence, more research is needed to determine its therapeutic value.

against strains of *M. tuberculosis* that are multidrug resistant to isoniazid and rifampin, which are therefore critical global public health problems. Its most common toxic effect is hepatotoxicity. It also interferes with the renal excretion of urate, resulting in hyperuricaemia (*Chapter 8*). Unfortunately, resistance to pyrazinamide is increasing following the evolution of a bacterial pyrazinamidase.

Drugs for treating eukaryotic pathogens and parasites

Protozoa, fungi and helminths are eukaryotic pathogens and parasites. They are responsible for many infections, particularly in developing countries. The drugs utilised to treat these infections may kill the causative organisms or simply inhibit their growth. In the latter case, therapy must be continued for sufficient time to allow the host immune system to eradicate the organism. Unfortunately, the drugs used in these treatments may also act against the patients' (eukaryotic) cells and so they often have undesirable side effects.

Antiprotozoal drugs

Protozoan diseases cause global health, social and economic problems. They are particularly prevalent in subtropical and tropical regions, largely because of the absence of suitable vaccines, difficulties in controlling their vectors (*Chapter 2*) and, unfortunately, the limited number of effective antiprotozoal drugs. Examples of antiprotozoal drugs are listed in *Table 3.6*, while their clinical target(s) and mechanisms of action of selected examples are outlined below.

Protozoal disease	Causative organism(s)	Drug(s) for treatment
African trypanosomiasis or sleeping sickness	*Trypanosoma brucei gambiense*, *T. brucei rhodesiense*	Eflornithine, melarsoprol, nifurtimox, pentamidine, suramin
Amoebiasis and amoebic dysentery	*Entamoeba histolytica*	Amphotericin B, dehydroemetine, metronidazole
Chagas disease or American trypanosomiasis	*Trypanosoma cruzi*	Benznidazole, nifurtimox
Giardiasis	*Giardia lamblia*, syn. *G. duodenalis*, *G. intestinalis*	Metronidazole, tinidazole
Leishmaniasis	*Leishmania* spp.	Amphotericin B, stibogluconate
Malaria	*Plasmodium* spp.	Artemisinin and its derivatives, artemether, artesunate, dihydroartemisinin, amodiaquine, lumefantrine, mefloquine, pentamidine, sulphadoxine-pyrimethamine, suramin, piperaquine
Toxoplasmosis	*Toxoplasma gondii*	Pentamidine, pyrimethamine with sulphonamide and folinic acid
Trichomoniasis	*Trichomonas vaginalis*	Metronidazole, pentamidine, tinidazole

Table 3.6 Examples of protozoan diseases and antiprotozoal drugs. ⏎

Artemisinin derivatives, that is, artemether, artesunate and dihydroartemisinin, are used in malarial therapy. They are activated by Fe^{2+}-containing proteins, including haems. These reactions produce highly reactive intermediates that are thought to alkylate and so inactivate essential parasite proteins, which results in the rapid death of susceptible species. These drugs are now usually used in combination therapies with other therapeutics, such as amodiaquine, mefloquine and piperaquine, which have alternative modes of action. For example, although its therapeutic mechanism has not been fully determined, it is thought that amodiaquine (see *Box 2.4*) and other quinoline derivatives inhibit haem polymerase activity, which results in an accumulation of free haem to concentrations that are toxic to the parasites.

Dehydroemetine is used in the treatment of amoebiasis. It is a synthetic derivative of emetine, which was first isolated from roots of *Cephaelis ipecacuanha* in 1828. Dehydroemetine binds to the 40S ribosomal subunit and therefore is thought to inhibit mRNA translation by preventing translocation of the tRNA–amino acid complex. Dehydroemetine is an effective amoebicide, although it has no effect on the cysts of *E. histolytica*.

Metronidazole is used to treat several different protozoan parasites, as well as bacterial pathogens (see above) that colonise anaerobic sites in the patient. It is reduced metabolically by ferredoxin or flavodoxin, generating short-lived free radicals that are cytotoxic because they bind to DNA and block nucleic acid synthesis.

Nifurtimox is used in to treat Chagas disease and African trypanosomiasis. It inhibits trypanothione reductase (TR), an enzyme that is both vital for the survival of the parasites and specific to them. Trypanothione and TR are unique to trypanosomes that function in a manner like that of the glutathione-glutathione reductase system described in *Chapter 13*. Thus, the reactions catalysed by these enzymes decrease the formation of free radicals. Inhibition of TR by nifurtimox (*Figure 3.54*) therefore induces oxidative stress in the parasites.

Pentamidine or suramin is recommended for treating the first stage of trypanosome infections. However, its second stage is treated with a combination of eflornithine and nifurtimox, both of which can cross the blood–brain barrier (see *Figure 3.8*) to target the parasites residing in the CNS. The mode of action of pentamidine is not well understood. It is thought to interfere with DNA, RNA, protein and phospholipid biosyntheses. It has been reported to inhibit DNA topoisomerases and also affects mitochondria, disrupting their inner membrane potential and inducing apoptotic death (see *Box 17.3*).

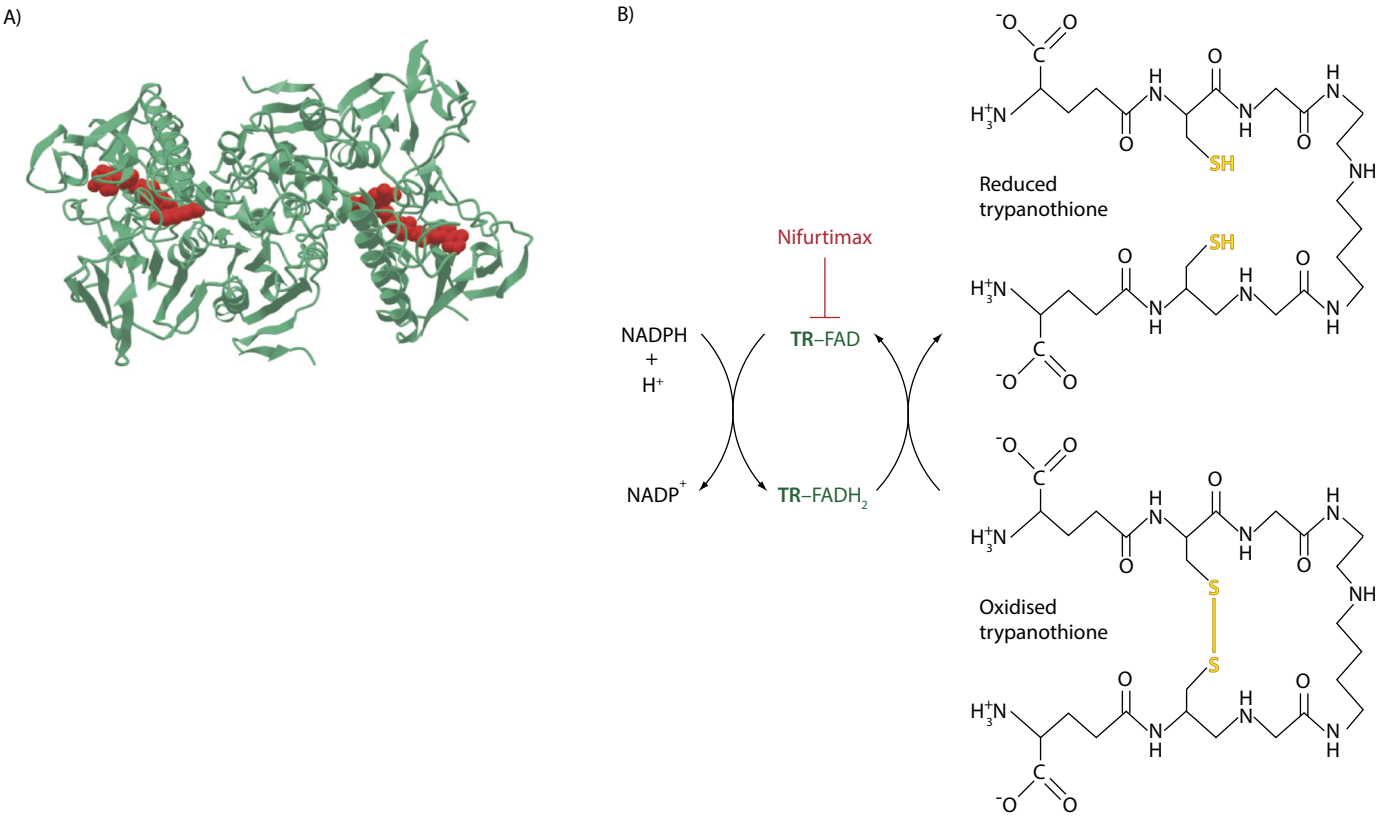


Figure 3.54 (A) Molecular model of the trypanothione reductase (TR) of *Trypanosoma cruzi*, the causative organism of Chagas disease. Note the two FAD coenzymes indicated in red. PDB file 1AOG. (B) The reaction catalysed by TR and inhibited by nifurtimox. Compare to *Figure 13.41*.

Suramin combines with proteins, especially the glycolytic enzymes of the trypanosome, and so inhibits energy metabolism. However, it is also a competitive inhibitor of gluconate-6-phosphate dehydrogenase, an enzyme of the pentose phosphate pathway (*Figure 3.55*). Its inhibition reduces supplies of NADPH, which is essential for reductive biosynthesis, and the pentose sugars necessary to biosynthesise nucleotides and nucleic

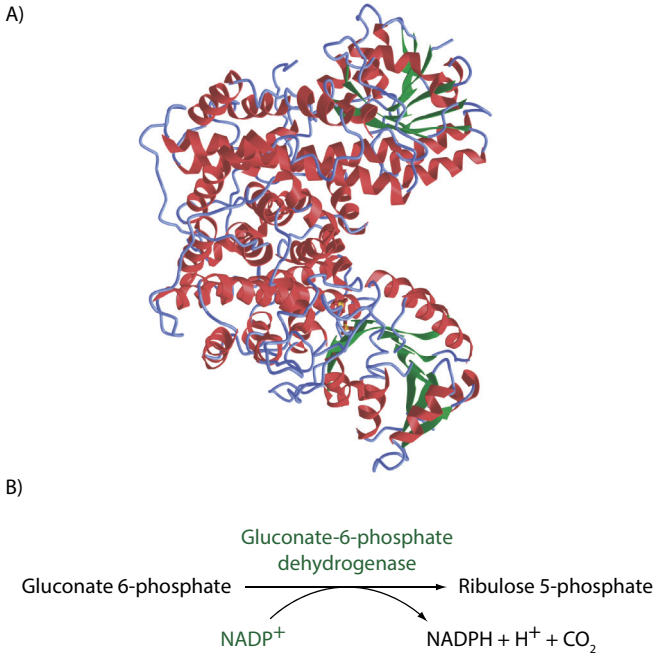


Figure 3.55 (A) Molecular model of the gluconate-6-phosphate dehydrogenase of *Trypanosoma brucei*. PDB file 1PGJ. (B) The reaction of the pentose phosphate pathway catalysed by gluconate-6-phosphate dehydrogenase.

acids. It has been reported to inhibit cytokinesis in the parasites. Melarsoprol can also cross the blood–brain barrier (see *Figure 3.8*) and was used as a monotherapeutic but, because of its high toxicity, it has largely been replaced with eflornithine. Eflornithine is used to treat second-stage cases of African sleeping sickness. It acts by irreversibly inhibiting ornithine decarboxylase, an enzyme that converts ornithine to the polyamine putrescine (*Figure 3.56*). Fexinidazole is also used in the treatment of late-stage *T. brucei gambiense* infections. It is metabolised *in vivo* by several cytochrome P-450s to two biologically active metabolites, fexinidazole sulphoxide and fexinidazole sulphone (*Figure 3.57*), which damage macromolecules in susceptible parasites.

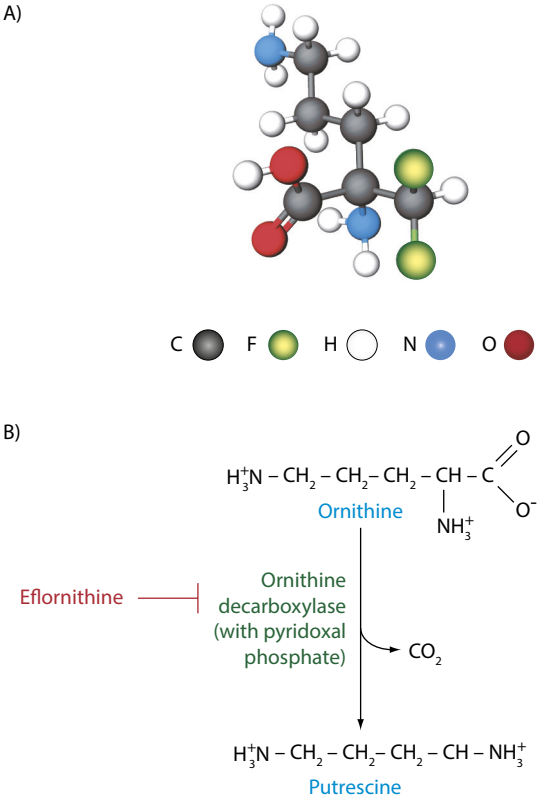


Figure 3.56 (A) Computer drawn model of eflornithine, which (B) inhibits the reaction catalysed by ornithine decarboxylase.

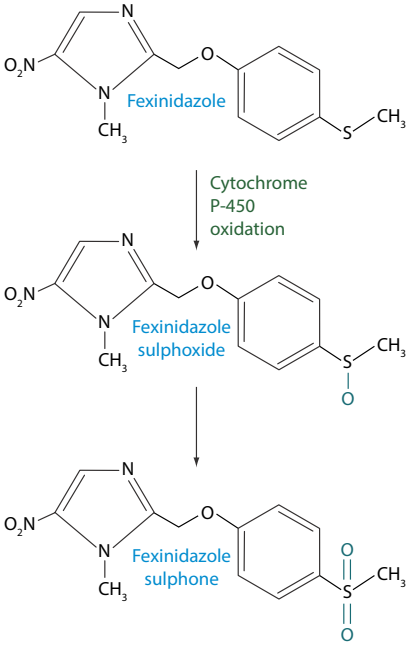


Figure 3.57 The conversion of fexinidazole to fexinidazole sulphoxide and fexinidazole sulphone.

Stibogluconate (*Figure 3.58*) is used to treat all forms of leishmaniasis. Its mechanism of action is poorly understood, but is thought to inhibit glycolysis and the TCA cycle, leading to reduced biosyntheses of ATP and GTP. However, it has been suggested that the major effect of this decrease is a reduction in the biosynthesis of parasite macromolecules that, in turn, decreases parasite viability. Other agents used in the treatment of leishmaniasis, and some forms of amoebiasis, include amphotericin B. This drug is also an antifungal agent (see below). Amphotericin B interacts with plasma membrane sterols and increases the permeability of the membrane to monovalent cations, such as H^+, Na^+ and K^+.

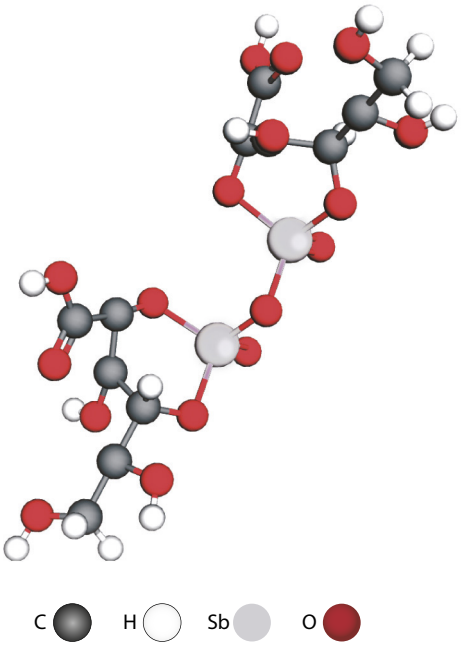


Figure 3.58 Computer drawn model of stibogluconate.

Tinidazole resembles metronidazole in structure and action (*Figure 3.59*). It also is a prodrug that is reduced by ferredoxin in *T. vaginalis* with the production of free nitro radicals. These toxic compounds are thought to be responsible for the antiprotozoal activity by covalently binding to DNA and the subsequent damage to it causing death of the parasite. Although the mechanism(s) of action against *G. lamblia* and *E. histolytica* are not known, they are likely to be similar.

Two major drawbacks are associated with antiprotozoal drugs. Their treatment can give rise to unwanted side effects. Gastrointestinal irritation, skin rashes, headaches, dizziness and nausea are common; severe toxic problems can involve eye disorders and kidney damage. Sadly, the widespread emergence of antimicrobial-resistant (AMR) parasites is reducing the effectiveness of current antiprotozoal therapies.

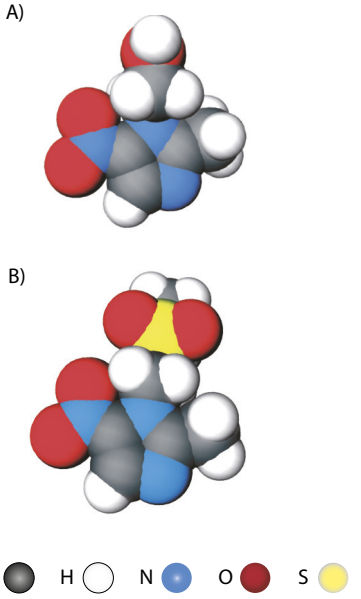


Figure 3.59 Computer drawn models of (A) metronidazole and (B) tinidazole.

Antifungal drugs

Antifungal drugs can be grouped into five major classes:

1. allylamines and benzylamines
2. echinocandins
3. flucytosine
4. polyenes
5. triazoles

Antifungal allylamines, such as terbinafine and naftifine, and the benzylamine butenafine are used to treat dermatophytosis caused by *Microsporum*, *Trichophyton* and *Epidermophyton* spp. (*Chapter 2*). Both allylamines and benzylamine are used topically (although terbinafine can also be given orally). Both types also act in a similar manner: they inhibit squalene epoxidase, an enzyme of the polyprenly pathway (*Chapter 5*), and so block the conversion of squalene to squalene epoxide (*Figure 3.60*). This action prevents the biosynthesis of ergosterol, an essential component of the fungal cell membrane. The decreased ergosterol in the membrane increases its general permeability and allows cytoplasmic contents to leak from the cell, while ions and water can enter freely.

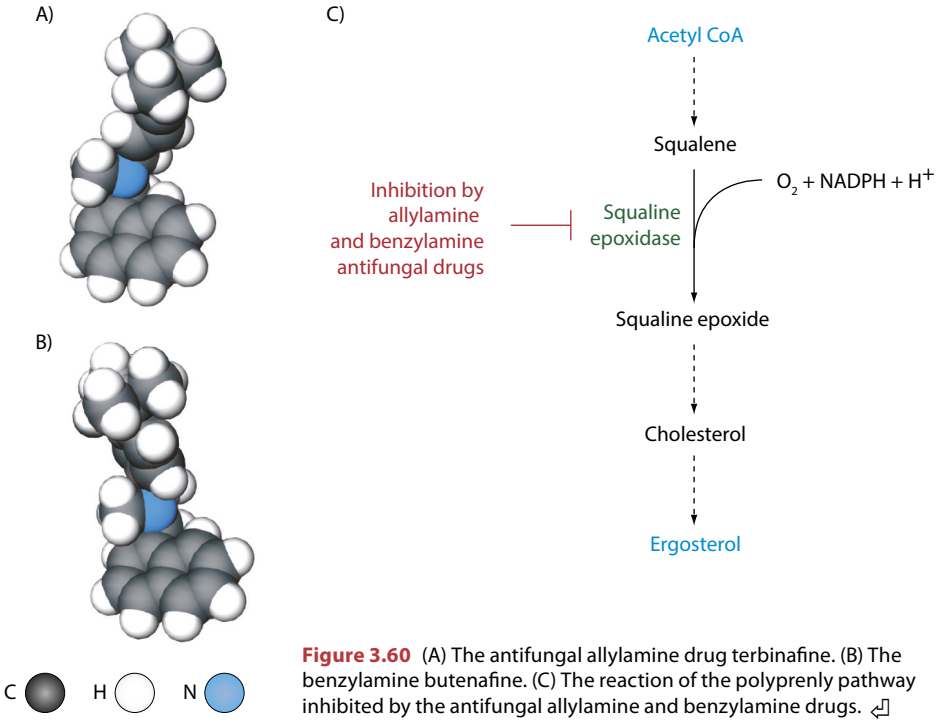


Figure 3.60 (A) The antifungal allylamine drug terbinafine. (B) The benzylamine butenafine. (C) The reaction of the polyprenly pathway inhibited by the antifungal allylamine and benzylamine drugs.

Echinocandins, for example micafungin and caspofungin, are the most important of the antifungal drugs because of their broad activities and less serious sides effects than other therapies. Indeed, they are the first-line treatment for patients suffering from invasive *Candida* infections. Echinocandins inhibit UDP-glucose β,1→3 glucosyltransferase (commonly referred as glucan synthase):

$$\text{UDP-glucose+Glucan}_n \xrightarrow{\text{Glucan synthase}} \text{UDP+Glucan}_{n+1}$$

This impairs the biosynthesis of the β,1→3 glucan polysaccharide, which is an essential component of the cell walls of many fungi.

5-Flucytosine (see *Margin Note 3.5*), a synthetic pyrimidine analogue, is used in combination with amphotericin B (see below) to treat severe systemic mycoses, for example aspergillosis, life-threatening *Candida* infections and cryptococcal meningitis. It is a prodrug and, following its absorption by susceptible fungal cells, is converted to the antimycotic 5-fluorouracil, which causes early chain termination during the biosynthesis of RNA and blocks DNA replication (*Figure 3.61*).

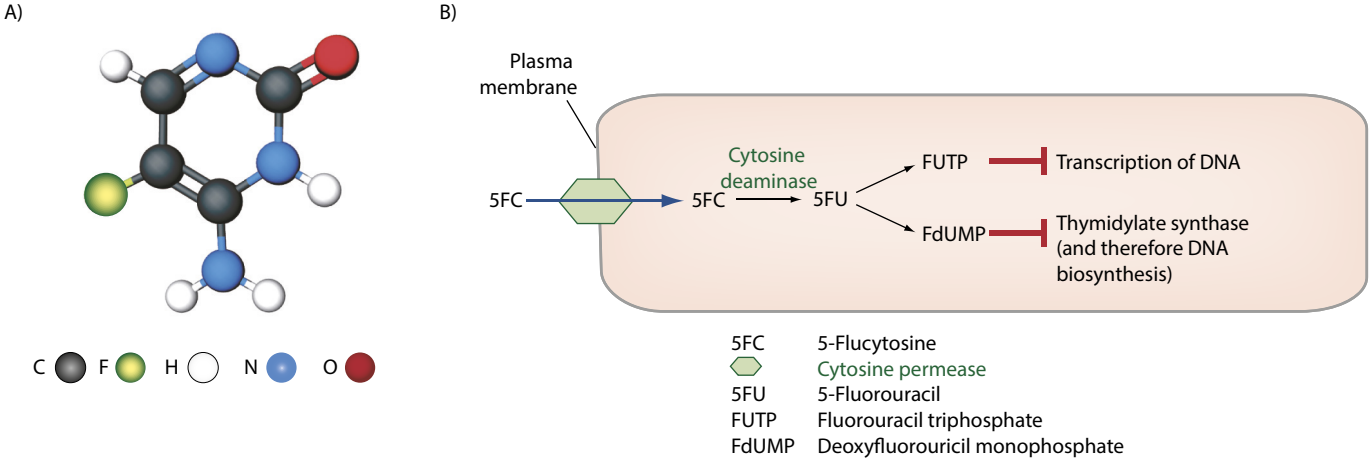


Figure 3.61 (A) Computer drawn model of 5-flucytosine. (B) Schematic illustration of the action of 5-flucytosine as an antifungal drug.

Amphotericin B deoxycholate, more usually referred to simply as amphotericin B (*Figure 3.62*), is perhaps the best-known polyene antifungal. It has been used for over 50 years to treat serious systemic fungal infections including invasive aspergillosis,

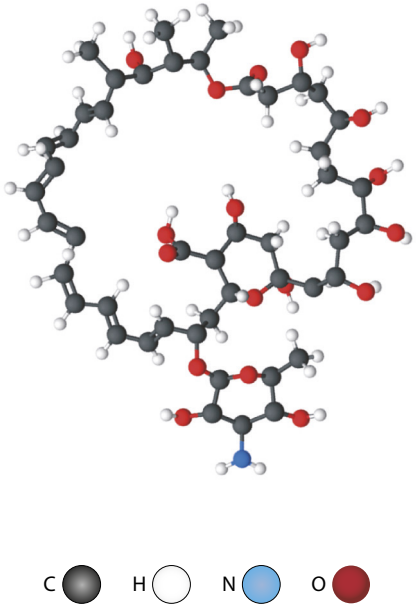


Figure 3.62 Computer drawn model of amphotericin B.

candidiasis, coccidioidomycosis, cryptococcosis (which affects the lungs and brain: cryptococcal meningitis) and histoplasmosis (*Chapter 2*). Amphotericin B is often used therapeutically in combination with other antifungals. Its therapeutic action is dependent upon its property of binding to ergosterol in the plasma membranes of fungal cells, causing them to aggregate and form a transmembrane pore. The resulting increase in cell membrane permeability allows cytoplasmic contents to leak and ions and water to enter, resulting in the death of the cell.

Triazoles are heterocyclic compounds with a broad range of clinical applications. Some are anticancer agents (*Chapter 17*); others are antiviral in action; examples of antifungal triazole drugs include fluconazole and voriconazole. These drugs function by inhibiting the fungal cytochrome P-450 14α-demethylase (also called CYP51; *Figure 3.63*) but are less effective against human P-450 enzymes (*Chapter 12*), an action that reduces the production of ergosterol necessary for fungal membranes.

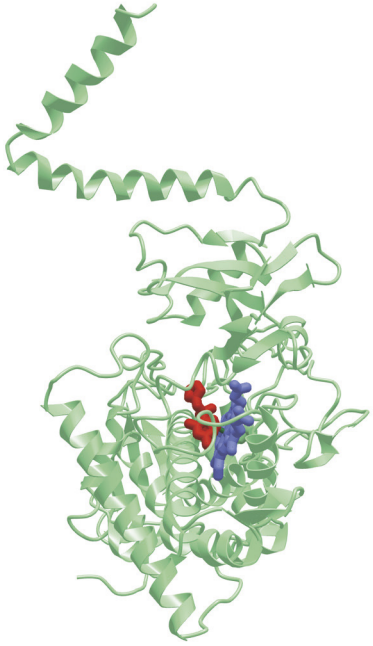


Figure 3.63 Molecular model of a cytochrome P-450 14α-demethylase of *Saccharomyces cerevisiae* with a fluconazole (red) bound to its haem group (blue). PDB file 5ESJ.

Toxic side effects are sometimes associated with the use of all antifungals. They appear to be least marked with echinocandins; however, those with amphotericin B can sometimes be so severe that therapy must be discontinued, even in patients with

life-threatening systemic fungal infections. Common adverse reactions to antifungal drugs include dizziness and headaches, itchiness, irritation and reddening of the skin because of dilated blood vessels, nausea and vomiting. More serious effects include fevers, hepatitis, thrombophlebitis and nephrotoxicity.

Fungi are no exception to the rule that MDR evolves in previously susceptible parasites. Fungi that are multidrug resistant, especially to amphotericin B, 5-flucytosine and triazole drugs, are increasingly reported to the detriment of patients.

Margin Note 3.8
New antifungal drugs

The research underpinning the development of new antifungal therapies has been sadly neglected, despite the severity of invasive fungal infection and the seriousness of their possible outcomes. However, the situation is now looking less bleak, as, from 2022 onwards, a number of new antifungal drugs, for example fosmanogepix, ibrexafungerp, rezafungin and olorofim, are in advanced trials or being approved by regulatory bodies such as the United States Food and Drug Administration (FDA).

Anthelmintic drugs

Anthelmintic, sometimes called antihelminthic, drugs can be vermifuges or vermicides. The former stimulate peristalsis to such a degree that the parasite is expelled from the GIT. The latter cause metabolic dysfunction or induce paralysis that kills the worm or increases its susceptibility to immune attack (*Chapter 4*). Anthelmintic drugs are often classified into cestodicides, nematicides and trematocides, based on their target group (*Figure 3.64*), although there is an overlap with some of these agents (*Table 3.7*). Given that the life cycles and metabolisms of helminths vary considerably between the different groups, it is necessary that the drugs have a range of modes of action.

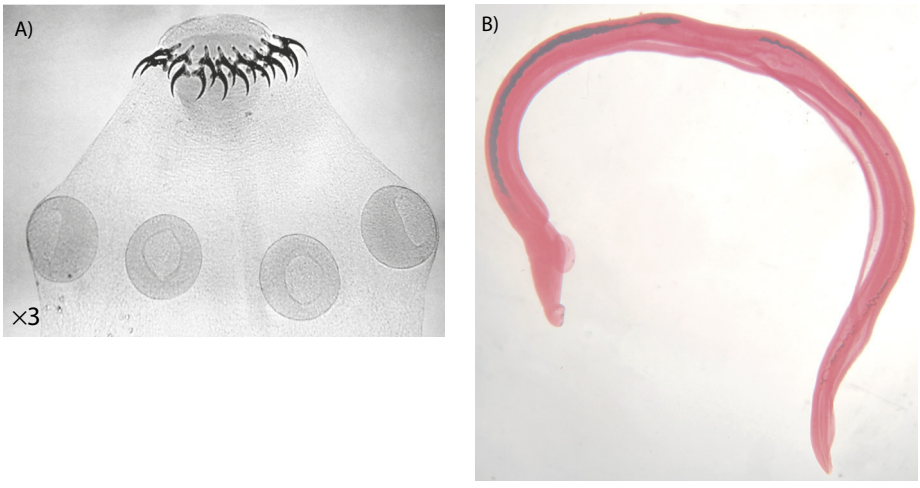


Figure 3.64 Examples of parasitic helminths ('worms'). (A) The 'head' of the pork tapeworm (a cestode) *Taenia solium*. Note the hooks and suckers that allow the intestinal parasite to remain attached to the gut wall. (B) *Schistosoma* sp. (a trematode), the causative organism of schistosomiasis/bilharzia. (A), Courtesy of the Centers for Disease Control and Prevention; (B), courtesy of Shutterstock.com/Mohammed Al Ali. See also Figure 2.42.

Helminth type	Examples of helminth parasite	Examples of anthelmintic drugs
Cestodes	Tapeworms, for example *Taenia solium*[a]	Praziquantel
Nematodes	Roundworms, for example *Ascaris lumbricoides*[a];	Albendazole, mebendazole
	Pinworms, for example *Enterococcus vermicularis*; threadworms, for example *Strongyloides stercoralis*	Albendazole, diethylcarbamazine, ivermectin, mebendazole, pyrantel pamoate
	Filarial worms, for example *Wuchereria bancrofti,*	Diethylcarbamazine, ivermectin
	Trichinella spiralis	Albendazole
	Hookworms, for example *Necator americanus*	Albendazole, diethylcarbamazine, mebendazole
Trematodes	Flukes, for example *Schistosoma* spp.[a]	Praziquantel

[a] See *Figure 2.42*.

Table 3.7 Examples of anthelmintic drugs.

The benzimidazole anthelmintics, such as albendazole and mebendazole (*Figure 3.65*), immobilise worms by binding to β tubulin monomers with an affinity that inhibits their polymerisation, disrupting the formation and necessary rearranging of microtubules of the cytoskeleton (*Chapter 16*). This prevents the microtubule-dependent uptake of glucose, decreasing the production of ATP and disrupting the energy metabolism of susceptible adult helminths living in the GIT, in addition to disrupting cell motility and reproduction. These agents are approved for treating GIT and tissue helminths; however, mebendazole is less effective than albendazole against internal helminths and is therefore used almost exclusively for intestinal nematode infections.

Pyrantel pamoate activates nicotinic cholinergic receptors on the surfaces of nematode somatic muscles, resulting in a persistent depolarising of the tissues. This produces an irreversible spastic paralysis in the worms, which lose their ability to adhere to the mucosal wall and are expelled on defecation and occasionally on vomiting.

Diethylcarbamazine (*Figure 3.66*) is the major treatment for lymphatic filariasis, since it leads to the rapid destruction of the microfilariae of, for example, *Wuchereria bancrofti* (*Chapter 2*) but also affects the muscular activities of adult worms, causing a hyperpolarisation that allows them to be dislodged. Diethylcarbamazine has been used clinically for over 40 years, but its full mechanism of action is still only partially understood. It inhibits lipoxygenase activity and so alters prostaglandin metabolism (*Chapter 10*) in microfilariae and host endothelial cells, leaving the microfilariae more susceptible to attack by the immune system. It is also thought to morphologically damage the membranes of microfilariae, so they are more susceptible to phagocytosis by monocytes.

Ivermectin acts by a different mechanism again. It binds selectively to glutamate gated Cl^- channels of invertebrate muscle and nerve cells. This increases their permeabilities to Cl^- and results in the hyperpolarisation of the cell, causing paralysis and death of the parasite. Ivermectin is also thought to be an agonist of the neurotransmitter γ aminobutyric acid (GABA) and so disrupts GABA-mediated nerve transmission.

Praziquantel (*Figure 3.67*), the drug of choice to treat schistosomiasis, is also active against larval and adult tapeworms. Despite being used for over 40 years to treat millions of patients, its mode of action is still the subject of research. Among several suggested mechanisms, it does appear to bind to the voltage sensor of voltage-gated Ca^{2+} channels, disrupting their regulation and increasing the permeability of the membranes to Ca^{2+}. The uncontrolled influx of Ca^{2+} into the cells stimulates uncontrolled contractions of the muscles, producing a paralysis that allows the worm to be eliminated.

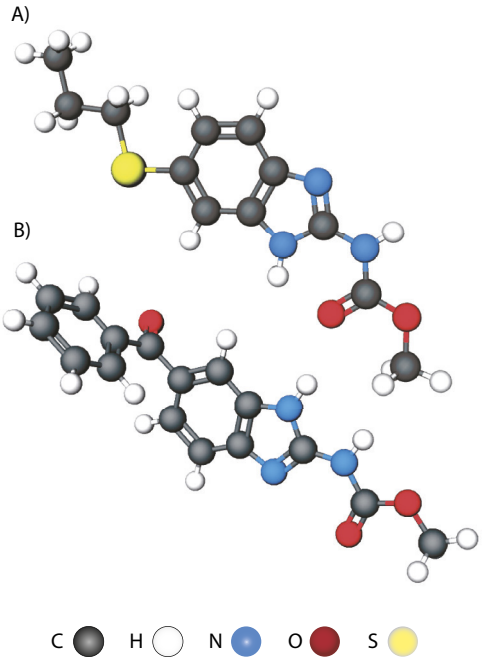


Figure 3.65 Computer drawn models of (A) albendazole and (B) mebendazole. ⏎

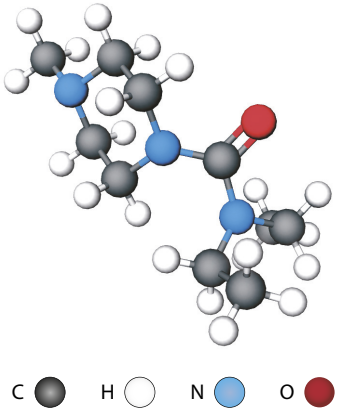


Figure 3.66 Computer drawn model of diethylcarbamazine. ⏎

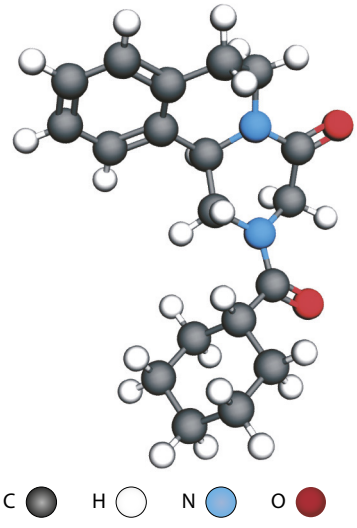


Figure 3.67 Computer drawn model of praziquantel. ⏎

As with other pathogen and parasite drugs, resistance among worms is now widespread. To combat this, anthelmintic drugs are sometimes used in combinations; for example, albendazole is combined with diethylcarbamazine or ivermectin for some conditions.

Many, indeed most, anthelmintic drugs produce unwanted side effects, including abdominal pain, dizziness and headaches, loss of appetite, nausea and vomiting, and weakness and fatigue.

Surgery

Most infectious diseases are treated using drug therapies. However, surgical intervention may be required in instances when the pathogen is resistant to available treatments or it is the only means to contain an infection that is spreading to other areas of the body, such as in gangrene caused by *Clostridium perfringens* (*Figure 3.68*). In other cases, surgery may be desirable because access to the affected site by the antimicrobial agents is limited, as in the case of some abscesses or in appendectomy where surgical drainage or removal of necrotic tissue, respectively, can enhance the recovery process.

Necrotising fasciitis is an infection with the so-called flesh-eating bacteria, for example group A streptococci, most commonly *S. pyogenes*. The prevalence and incidence of this condition are both rare but can begin after surgery, particularly abdominal or gynaecological interventions. However, it may also develop from complications of childbirth, burns or following relatively minor traumas, for example bites and abscesses. The bacteria produce extracellular enzymes that attack soft tissues, often of extremities, causing an extensive necrosis of subcutaneous tissues. The lysosomal hydrolytic enzymes (*Chapter 16*) released from damaged cells of the patient may exacerbate the bacterial damage. The condition can be diagnosed by culturing the bacteria from blood samples or aspirations of pus from affected sites, although surgical exploration may be necessary. Necrotising fasciitis can have such a sudden and rapid onset that the extent of soft-tissue destruction may be fatal. Hence, an early diagnosis and prompt medical and surgical interventions are necessary to reduce the risk of death. Treatment often includes intravenous penicillin and clindamycin, along with aggressive surgical **debridement**, that is, the removal of infected tissue, which can be very extensive. For those with severe illness, admission to an intensive care unit is needed. Limb amputation may be necessary. Unfortunately, approximately 20% of necrotising fasciitis patients die of the condition.

Surgery may also be necessary to treat some eukaryotic infections, especially in immunocompromised patients. For example, surgery may be necessary to treat invasive sinusitis and pulmonary aspergillosis, when chemotherapy has failed to control the infection. Other fungi that may be involved include *Candida* and *Fusarium* spp. Surgical therapy has also been used to treat infections of the CNS, particularly with the fungi *Coccidioides immitis* and *Cryptococcus neoformans*. Fungal nail infections that fail to respond to treatment may require the removal of the nail. Helminth infestation may also require surgical intervention. Intestinal obstruction by *Ascaris lumbricoides* in particular may necessitate emergency surgery (*Figure 3.69*). Cysts of the tapeworm *E. granulosus* in patients with hydatid disease are also removed in surgical procedures.

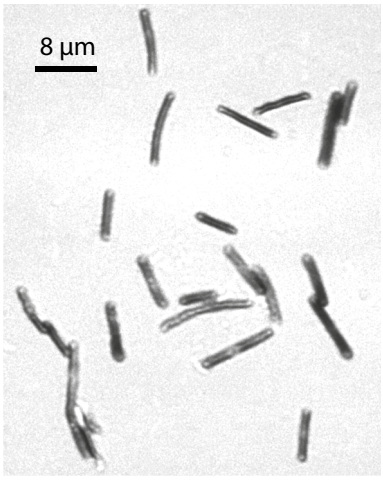


Figure 3.68 Photomicrograph of *Clostridium perfringens* growing in Schaedler's broth. Courtesy of the Centers for Disease Control and Prevention/D. Stalons. ⏎

Figure 3.69 The surgically removed section of intestine from a three-year-old boy at Red Cross Children's Hospital, Western Cape, South Africa. The operation was essential because of a blockage by a large infestation of *Ascaris lumbricoides*. Courtesy of Allen Jefthas and the South African Medical Research Council. ⏎

3.12 CASE STUDIES

CASE STUDY 3.1

Ellie, a five-year-old girl, presented to her general practitioner generally feeling unwell, with fever, joint pains, loss of appetite and diffuse itchy lesions particularly over her head, face, back and chest and even the inside of her mouth. These lesions started about three to four days prior and had begun as numerous small red pimples that developed into fluid-filled blisters, many of which were progressing to form scabs. Her parents belong to a religious group that does not approve of medical interventions, such as vaccinations and blood transfusions.

Questions

(a) From what is Ellie likely to be suffering?
(b) How should Ellie be treated?
(c) What could have been done to protect Ellie from this condition?
(d) Is Ellie likely to get this condition again?

CASE STUDY 3.2

Andrew, a 48-year-old male, presented at the accident and emergency department of his local hospital. He told the examining doctor that the previous day he had experienced pain on the side of his right calf. Within six hours, the area was septic and pus could be expressed. Over the day, the swelling and redness (inflammation) gradually extended to cover the area from the ankle to the knee. Physical examination showed respiration, pulse, blood pressure and temperature were all within reference values. There was no obvious area of pus at the site and pus did not drain on puncturing with a 20-gauge needle. Andrew was treated with intramuscular and oral ceftriaxone. Two days later, Andrew returned with increased pain and slightly increased temperature. The centre of inflammation showed an obvious pus-filled area. Culture of the pus showed the presence of clusters of Gram-positive cocci on nutrient agar and yellowish-coloured colonies on 5% sheep blood agar. The pus was removed by aspiration and the area excised and drained.

Questions

(a) Describe the history of this case.
(b) Suggest why antibiotic treatment alone was insufficient and Andrew required drainage and excision at the site of infection.

CASE STUDY 3.3

Chris, a six-year-old boy, complained of a sore throat and had a temperature of 38.6°C. His mother kept him home from school and treated him with a proprietary paediatric painkiller. He slept well but awoke with similar symptoms. His mother took him to their family doctor. Physical examination showed reddening of the nasopharyngeal area and tonsils and a slight enlargement of the cervical lymph nodes. The skin was clear.

Questions

(a) What is/are the most likely organism(s) causing this disease?
(b) How could these organisms be detected?
(c) What treatment is desirable and why?

CASE STUDY 3.4

Neha, a six-week-old baby, was transferred to hospital after several days of recurring attacks of repeated severe coughing that left her gasping for breath and that were eventually associated with bouts of vomiting. Physical examination showed a normal trachea, and a radiograph showed the chest to be clear of infection. Clinical examination showed the following data (reference values in parentheses):

- pulse of 155/min (100–160/min for babies younger than three months old)
- respiratory rate of 71/min (30–60/min for babies and children)
- leukocyte count of 15 000 000/mL (5 000 000 to 21 000 000/mL for babies two to eight weeks old)

A nasopharyngeal swab was taken. Organisms did not culture on standard media. Bacterial plaques with a mercury-like appearance were visible when charcoal blood agar was used as culture medium.

Questions

(a) What is Neha suffering from?
(b) What is the causative organism?

CASE STUDY 3.5

Jim, a 50-year-old male, consulted his ophthalmologist for some loss of vision in his left eye. He had previously reported a blind spot in the centre of this eye but subsequently his vision had deteriorated and he reported having regular headaches during this period. Following an angiogram, the ophthalmologist found evidence for optic nerve oedema. Several investigations were conducted including a magnetic resonance imaging (MRI) scan of the brain and tests for toxoplasmosis and sarcoidosis. The results of these investigations were generally inconclusive except that Jim tested positive for immunoglobulin G (IgG) antibodies against toxoplasma.

Questions

(a) What might be responsible for Jim's loss of vision?
(b) How should Jim be treated?
(c) What is the likely prognosis regarding Jim's vision?

3.13 SUMMARY

Pathogenic viruses, bacteria and eukaryotic parasites cause clinical conditions ranging from infections of single organs to systemic disease. Infections of the skin may be caused by viruses, such as papilloma and herpes simplex types. Skin bacteria cause boils and abscesses and severe and chronic conditions such as leprosy. Fungi also cause a variety of skin infections, although these are not usually severe, unless the patient is immunocompromised. Severe infections of the eye include trachoma, a chlamydial disease that is the most common eye infection worldwide. Helminths, such as *T. canis*, can cause blindness. Viruses and bacteria may cause painful ear infections, which, if recurrent, may damage hearing. Both groups of organisms can infect the CNS, sometimes resulting in encephalomyelitis by herpes simplex or meningitis by *N. meningitides*. Several viruses and bacteria cause significant respiratory infections, some of which, for example TB, are chronic. Many species of pathogens and parasites can infect the GIT, causing illnesses ranging from short-term gastroenteritis to stomach ulcers. An extensive number of microorganisms can infect the genital tract or are transmitted sexually. Several may cause systemic diseases.

The diagnosis of infectious disease involves a combination of laboratory tests to identify the organism(s) present in specimens, typically by culturing them. However, the presence of antibodies to the organisms in the patient may be used for identification and to compare levels in acute and convalescent patients. An increasing number of biochemical and molecular biological tests are also available, many of which rely on identifying viral or bacterial nucleic acid in clinical specimens. Imaging techniques are increasingly used for diagnostic purposes.

Infectious diseases may be prevented by public health measures and vaccination or prophylactic treatment with drugs, which are also used to treat infections. Depending on the infectious agent, these may be antiviral, antibacterial (antibiotics), antiprotozoal, antifungal or anthelmintic. Drugs may be given singly or in combinations, such as in the treatment of HIV patients. In immunocompromised patients particularly, surgical interventions may be necessary to control some types of infections.

QUESTIONS

1. Briefly review the causes of infections associated with the genital tract.
2. List some of the methods used in the diagnosis of infectious diseases.
3. A 19-year-old student attended a barbecue to celebrate his brother's 21st birthday. He ate a very rare steak. Two days later, he began to experience pains on defecation and to suffer from persistent diarrhoea that rapidly increased in frequency and produced bloody stools. What is the student suffering from? Suggest how this condition should be treated.
4. Examine *Figure 3.33*. To which antibiotics is this strain of *E. coli* resistant?
5. In what ways may antibiotics prevent the growth of bacterial pathogens?
6. Compare and contrast the actions of griseofulvin and albendazole.
7. In two words, state a major way of delaying the development of antidrug resistance in pathogens.
8. Which of the following is the odd ONE out?
   (a) *Echinococcus granulosus*.
   (b) *Staphylococcus aureus*.
   (c) *Streptococcus sanguis*.
   (d) *Streptococcus mutans*.
   (e) *Haemophilus influenzae*.
9. Which of the following is the odd ONE out?
   (a) *Neisseria meningitidis*.
   (b) *Varicella zoster*.
   (c) *Toxoplasma gondii*.
   (d) *Microsporum* spp.
   (e) *Toxocara canis*.

10. Which of the following is (are) inappropriate pairings?

(a)	HHV-3	Chicken pox/shingles
(b)	HAART	HIV–AIDS
(c)	Phagotherapy	*Trichomonas vaginalis*
(d)	Rifamycins	80S ribosome
(e)	β Lactams	Cephalosporins
(f)	Sleeping sickness	*Trypanosoma brucei gambiense* and *T. brucei rhodesiense*
(g)	Allylamines	Dermatophytosis
(h)	Systemic mycoses	5-Flucytosine and amphotericin B
(i)	Toxoplasmosis	*Toxocara canis*
(j)	Surgery	Obstructive *Ascaris lumbricoides*

FURTHER READING

Abraham SN *et al.* (1998) Fimbriae-mediated host-pathogen cross-talk. *Curr Opin Microbiol* **1**:75–81.

Abuhammour W and Habte-Gaber E (2004) Newer antifungal agents. *Indian J Pediatr* **71**:253–259.

Adams RA *et al.* (2021) Rifamycin antibiotics and the mechanisms of their failure. *J Antibiot (Tokyo)* **74**:786–798.

Adamson CS *et al.* (2021) Antiviral drug discovery: preparing for the next pandemic. *Chem Soc Rev* **50**:3647–3655.

Ang CW *et al.* (2017) Nitroimidazoles: molecular fireworks that combat a broad spectrum of infectious diseases. *J Med Chem* **60**:7636–7657.

Ayoub Moubareck C (2020) Polymyxins and bacterial membranes: a review of antibacterial activity and mechanisms of resistance. *Membranes (Basel)* **10**:181. doi: 10.3390/membranes10080181.

Behja W and Jemal M (2019) Anti-HIV drug discovery, development and synthesis of delavirdine: review article. *IRJPAC* **20**:1–16.

Brives C and Pourraz J (2020) Phage therapy as a potential solution in the fight against AMR: obstacles and possible futures. *Palgrave Commun* **6**:100. doi: 10.1057/s41599-020-0478-4.

Bush NG *et al.* (2020) Quinolones: mechanism, lethality and their contributions to antibiotic resistance. *Molecules* **25**:5662. doi: 10.3390/molecules25235662.

Capela R *et al.* (2019) An overview of drug resistance in protozoal diseases. *Int J Mol Sci* **20**:5748. doi: 10.3390/ijms20225748.

Chakraborty C *et al.* (2005) Prion disease: a deadly disease for protein misfolding. *Curr Pharm Biotechnol* **6**:167–177.

Chaudhry SB *et al.* (2019) Cephalosporins: a focus on side chains and β-lactam cross-reactivity. *Pharmacy (Basel)* **7**:103. doi: 10.3390/pharmacy7030103.

De Clercq E and Li G (2016) Approved antiviral drugs over the past 50 years. *Clin Microbiol Rev* **29**:695–747.

Dingsdag SA and Hunter N (2018) Metronidazole: an update on metabolism, structure-cytotoxicity and resistance mechanisms. *J Antimicrob Chemother* **73**:265–279.

Dispinseri S *et al.* (2022) The path towards an HIV vaccine. *New Microbiol* **45**:99–103.

Dolgin E (2021) The tangled history of mRNA vaccines. *Nature* **597**:318–324.

Frieden TR *et al.* (2003) Tuberculosis. *Lancet* **362**:887–899.

Gaynes R (2017) The discovery of penicillin–new insights after more than 75 years of clinical use. *Emerg Infect Dis* **23**:849–853.

Ghosn J *et al.* (2018) HIV. *Lancet*, **392**:685–697.

Gould D (2004) Bacterial infections: antibiotics and decontamination. *Nurs Stand* **18**:38–42.

Green LH and Goldman E (2021) *Practical Handbook of Microbiology*, 4th edn. Taylor & Francis.

Harada LK *et al.* (2018) Biotechnological applications of bacteriophages: state of the art. *Microbiol Res* **212–213**:38–58.

Hayden FG (2006) Antiviral resistance in influenza viruses – implications for management and pandemic response. *N Engl J Med* **354**:785–788.

Kaisers S *et al.* (2021) A review: mechanism of action of antiviral drugs. *Int J Immunopathol Pharmacol* **35**:20587384211002621. doi: 10.1177/20587384211002621.

Krause KM *et al.* (2016) Aminoglycosides: an overview. *Cold Spring Harb Perspect Med* **6**:a027029. doi: 10.1101/cshperspect.a027029.

Krcmery VC (2005) Antifungal chemotherapeutics. *Med Princ Pract* **14**:125–135.

Lavanya R (2017) Sulphonamides: a pharmaceutical review. *Int J Pharm Sci* **6**:1–3.

Lenz KD *et al.* (2021) Macrolides: from toxins to therapeutics. *Toxins (Basel)* **13**:347.

Lewis K (2020) The science of antibiotic discovery. *Cell* **181**:29–45.

Lima LM *et al.* (2020) β-lactam antibiotics: an overview from a medicinal chemistry perspective. *Eur J Med Chem* **208**:112829. doi: 10.1016/j.ejmech.2020.112829.

Magner LN and Kim OJ (2017) *A History of Medicine*, 3rd edn. Taylor & Francis.

Mann S *et al.* (2021) A review of leishmaniasis: Current Knowledge and Future Directions *Curr Trop Med Rep* **8**:121–132.

Maslow MJ and Portal-Celhay C (2020) Rifamycins. In: *Mandell, Douglas and Bennett's Principles and Practice of Infectious Diseases*, 9th edn. Elsevier, pp. 337–349.

Matin MM *et al.* (2022) Triazoles and their derivatives: chemistry, synthesis, and therapeutic applications. *Front Mol Biosci* **9**:864286. doi: 10.3389/fmolb.2022.864286.

Millanao AR *et al.* (2021) Biological effects of quinolones: a family of broad-spectrum antimicrobial agents. *Molecules* **26**:7153. doi: 10.3390/molecules26237153.

Mohapatra SS *et al.* (2021) Polymyxins, the last-resort antibiotics: mode of action, resistance emergence, and potential solutions. *J Biosc* **46**:85. doi: 10.1007/s12038-021-00209-8.

Mroczyńska M and Brillowska-Dąbrowska A (2020) Review on current status of echinocandins use. *Antibiotics (Basel)* **9**:227. doi: 10.3390/antibiotics9050227.

Murray PR *et al.* (2020) *Medical Microbiology*, 9th edn. Elsevier.

Ovung A and Bhattacharyya J (2021) Sulfonamide drugs: structure, antibacterial property, toxicity, and biophysical interactions. *Biophys Rev* **13**:259–272.

Papageorgiou AC and Acharya KR (2000) Microbial superantigens: from structure to function. *Trends Microbiol* **8**:369–375.

Prinz C *et al*. (2003) *Helicobacter pylori* virulence factors and the host immune response: implications for therapeutic vaccination. *Trends Microbiol* **11**:134–138.
Prusiner SB (2016) *Madness and Memory: The Discovery of Prions – A New Biological Principle of Disease*. Yale University Press.
Putman M *et al*. (2000) Molecular properties of bacterial multidrug transporters. *Microbiol Mol Biol Rev* **64**:672–693.
Rhen M *et al*. (2003) The basis of persistent bacterial infections. *Trends Microbiol* **11**:80–86.
Sangster NC *et al*. (2018) Ten events that defined anthelmintic resistance research. *Trends Parasitol* **34**:553–563.
Schoen RT (2020) Lyme disease: diagnosis and treatment. *Curr Opin Rheumatol* **32**:247–254.
Seeger C and Mason WS (2000) Hepatitis B virus biology. *Microbiol Mol Biol Rev* **64**:51–68.
Sharma R *et al*. (2005) Antibacterial resistance: current problems and possible solutions. *Indian J Med Sci* **59**:120–129.
Shaw-Taylor L (2020) An introduction to the history of infectious diseases, epidemics and the early phases of the long-run decline in mortality. *Econ Hist Rev* **73**:E1–E19.
Singh N and Ghosh M (2020) Evolution of multidrug resistance superbugs, a healthcare nightmare – short review. *Eur J Mol Clin Med* **7**:4378–4382.
Smith MA (2005) Antibiotic resistance. *Nurs Clin North Am* **40**:63–75.
Spížek J and Řezanka T (2017) Lincosamides: chemical structure, biosynthesis, mechanism of action, resistance, and applications. *Biochem Pharmacol* **133**:20–28.
Stebbing JS *et al*. (2004) Where does HIV live? *N Engl J Med* **350**:1872–1880.
Stevens DL and Bryant AE (2023) Endemic, epidemic and pandemic infections: the roles of natural and acquired herd immunity. *Curr Opin Infect Dis* **36**:171–176.
Tevyashova AN (2021) Recent trends in synthesis of chloramphenicol new derivatives. *Antibiotics (Basel)*, **10**:3700. doi: 10.3390/antibiotics10040370.
van Seventer JM and Hochberg NS (2017) Principles of infectious diseases: transmission, diagnosis, prevention, and control. In: *International Encyclopedia of Public Health*. Elsevier, pp. 22–39.
Vázquez-Laslop N and Mankin AS (2018) How macrolide antibiotics work. *Trends Biochem Sci* **43**:668–684.
Wong KH *et al*. (2022) A review of post-treatment Lyme disease syndrome and chronic Lyme disease for the practicing immunologist. *Clin Rev Allerg Immunol* **62**:264–271.
Yang S and Rothman RE (2004) PCR-based diagnostics for infectious diseases: uses, limitations, and future applications in acute-care settings. *Lancet Infect Dis* **4**:337–348.
Yarchoan R and Uldrick TS (2018) HIV-associated cancers and related diseases. *N Engl J Med* **378**:1029–1041.

Useful websites

www.ecdc.europa.eu
www.cdc.gov
www.cjd.ed.ac.uk
www.who.int
https://www.who.int/news-room/questions-and-answers/item/neglected-tropical-diseases

Immune system

4

OBJECTIVES

After studying this chapter you should be able to:

- give an account of the roles of the immune system in maintaining health
- explain the difference between nonspecific and specific immune defences
- explain the roles of immunoglobulins in immune defence
- discuss the differing types of cytokines and outline their activities in coordinating immune response
- describe the cells of the immune system and outline their roles in immune defence
- describe the major features of inflammation and the acute phase response
- describe the major features of humoral and cell-mediated specific immunities
- outline the uses of antibodies in detecting and quantifying molecules of biomedical significance

4.1 INTRODUCTION

The **immune system** is a set of molecules, cells, tissues and organs that protect the body from diseases caused by viruses, microorganisms and multicellular parasites (*Chapters 2* and *3*). **Immunology** is the study of the structure and functions of the immune system. Unravelling the mechanisms of immune defence and applying its products to investigate and treat clinical conditions and diseases has revolutionised the biomedical sciences in a way that is comparable to the advances made in molecular biology in recent decades. This chapter describes major components of the immune system and explains how their actions produce immune responses against infectious agents. It will outline the roles of signalling molecules called **cytokines**, which allow components of the immune system to communicate with each other to produce coordinated responses to infections. Some of the uses of immunological techniques in the biomedical sciences will also be described.

4.2 TYPES OF IMMUNE DEFENCE

The immune system has to distinguish between the macromolecules and cells making up the body, that is, **self** from those that are foreign, or **nonself**. Immunological defences are usually classified as **nonspecific** and **specific**. Nonspecific defences constitute a first line of defence and are available immediately any foreign material, for example wood splinters, or proteins, viruses, bacteria or multicellular parasites enter the body. Nonspecific defences include responses such as **inflammation**, a rapid immediate response to tissue damage (*Figure 4.1*), and the **acute phase response**, which is a relatively rapid response to infection.

Figure 4.1 Note the redness and swelling on this badly inflamed foot caused by a skin allergy to the antibiotic cefaclor. Courtesy of Dr Harry Gouvas (https://independent.academia.edu/DrHarryGouvas). ⏎

DOI: 10.1201/9781003296836-4

In specific immunological defence, cells of the immune system recognise not just individual pathogens or parasites, but also particular proteins or glycoproteins found on their surfaces (*Figure 4.2*). This type of defence may take several days to become effective, depending on whether the individual has been previously exposed to that specific nonself or foreign material. Once activated, specific immunological defence results in long-lasting immunity. This type of immunity can be **humoral**, which involves the production of antibodies, and/or **cell-mediated**, involving cells that kill or recruit other cells to kill infected cells. Humoral immunity is effective against microorganisms that do not invade cells, which includes most bacteria and multicellular parasites (*Chapter 2*). Cell-mediated immunity is effective against intracellular parasites, including viruses and some groups of bacteria. However, both types of specific immunity are not mutually exclusive and are usually jointly activated on exposure to an infectious agent.

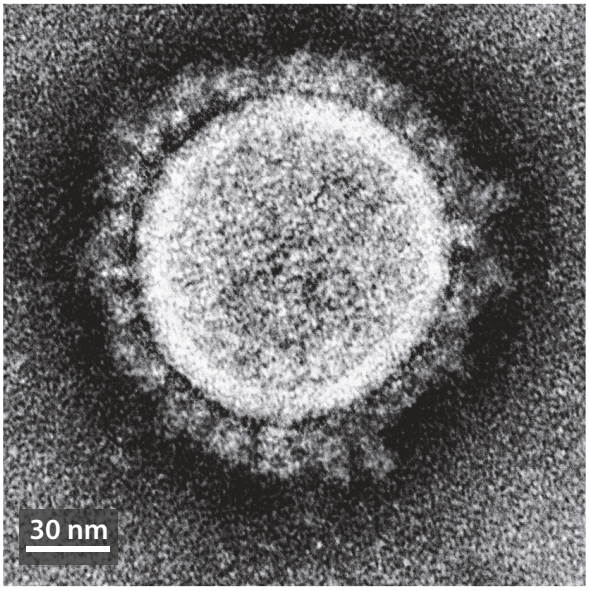


Figure 4.2 Electron micrograph of a severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) virion. Note the extended 'spike' glycoproteins that are antigenic and stimulate a specific immunological defence. Courtesy of the National Institute of Allergy and Infectious Diseases. ⏎

The activation, coordinated maintenance and cessation of immune responses requires strict physiological control and regulation: much of this is achieved by the appropriate secretions of signalling molecules called **cytokines** (see *Section 4.4*).

4.3 NONSPECIFIC DEFENCES

The entry of pathogens and parasites into the body is restricted by protective barriers, such as the skin and mucous membranes (see *Margin Note 4.1*). These structures offer further protection because of the (bio)chemical secretions they contain, for example lactic acid in sebum and hydrochloric acid in the stomach. Furthermore, mucus, secreted by the mucosal membranes that line the urogenital, respiratory and gastrointestinal tracts, contains lysozyme (*Figure 4.3*), an antibacterial enzyme that catalyses the hydrolysis of peptidoglycans in the cell walls of Gram-positive bacteria (*Chapter 2*). Should pathogens or parasites breach these barriers, other antimicrobial systems, such as complement and interferon proteins, may be activated and/or produced. Also, a number of cell types have roles in their elimination. These cells occur in the blood, although a number are also found in solid tissues. Some blood cells move between the blood and extravascular tissues, particularly during an infection (*Section 4.6*).

Margin Note 4.1
Mucous membranes and skin: more than just inert barriers ⏎

Mucous membranes are epithelial tissues that secrete a layer of mucus of varying thickness over their free, that is, apical, surface (see *Chapter 11*). Mucus consists of up to 80% water, held in place by glycoproteins, called mucins, and proteins, which constitute less than 20% of the mucus. In addition to offering structural protection, mucus also contains lysozyme and a variety of antibacterial peptides.

Skin has a number of well-documented physiological roles in addition to forming a protective barrier. These include contributing to thermal regulation, detecting environmental stimuli, excretion of relatively minor amounts of waste material and the synthesis of vitamin D (*Chapter 10*). Although skin forms a waterproof layer, it also has absorptive properties. Clinical use is made of this in a variety of topical treatments, for example acyclovir in antiviral gels to treat cold sores (*Chapters 2* and 3), phenol and cetrimide combined with chlorhexamide in proprietary antiseptic creams, and terbinafine ointments to treat fungal infections, such as athlete foot (*Chapter 3*), while 10% ibuprofen is used to reduce the inflammation associated with chronic lower back pain and moderate arthritic complaints.

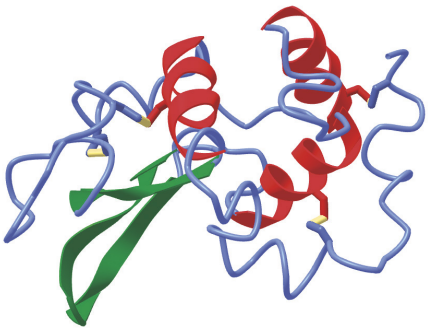


Figure 4.3 Molecular model of a lysozyme. PDB file 1REX.

4.4 CYTOKINES

Cytokines are proteins or glycoproteins of relatively low M_r (range 5000–70 000) that have signalling or intercellular communication roles that coordinate intercellular activities, especially, but not exclusively, those involved in immune responses. They are produced by a variety of cell types, including endothelial cells, fibroblasts, keratinocytes, lymphocytes, macrophages, mast cells and **natural killer cells** (NKCs). Cytokines stimulate activities in target cells by binding to specific receptors (*Figure 4.4*) on their surfaces. This stimulates the production of molecules, called **second messengers** (the cytokine being the first messenger), within the target cell. The increased intracellular concentration of the second messenger simulates a variety of activities, which involves altering the genes expressed by the cell. These actions can be autocrine by binding of the cytokine to receptors on the membrane of the cytokine-producing cell; paracrine, that is, acting locally; or, more rarely, endocrine, where distant cells are stimulated. Many cytokines act in a pleiotropic manner, with an individual cytokine producing multiple effects in one cell or in several different types of cells. Their actions can also be redundant, in that a number of different cytokines may induce similar physiological activities. Cytokines often act in complex, synergistic networks, with one enhancing the activities/production of others, while in some instances their effects may be antagonistic!

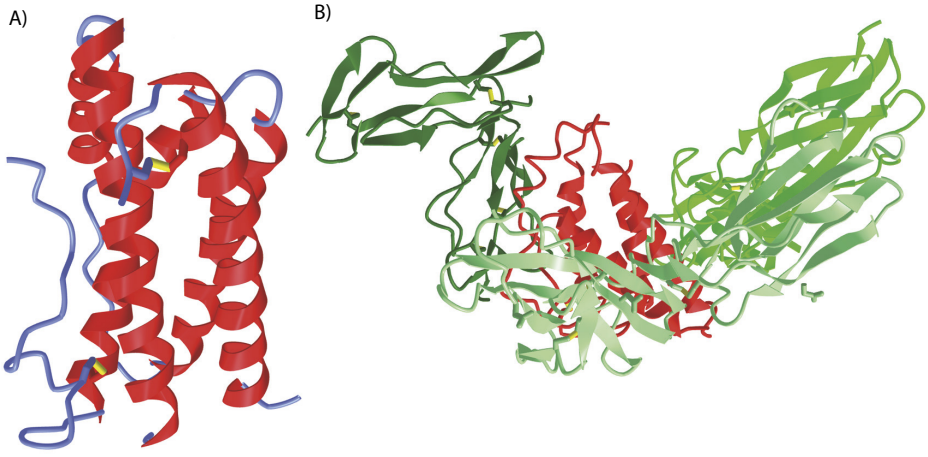


Figure 4.4 Molecular models showing the cytokines (A) interleukin-6 (IL-6) and (B) interleukin-2 (IL-2; red) bound to the α, β and γ subunits of its receptor, which are shown in dark, medium and light greens, respectively. PDB files 1ALU and 2B5I, respectively.

Cytokine nomenclature, classification and selected activities

Cytokine nomenclature is something of a mess. Given it was initially thought that they were produced by leukocytes, those produced by monocytes and lymphocytes were originally termed monokines and lymphokines, respectively. These terms were replaced by **interleukin** (IL) and cytokines numbered in order of their discovery. Unfortunately, a number of them were named before this attempted rationalisation and these original names have tended to be retained, for example tumour necrosis factor (TNF) and B cell activating factor (BAFF). The former cytokine was discovered in 1975 and named following the observation that it promoted the *in vitro* necrosis of tumours. B cell survival factor

(*Figure 4.5*) is a member of the TNF superfamily, which binds to receptors on mature B cells, stimulating the activation of transcription factors, leading to the survival of the cell and secretion of antibodies (*Section 4.8*).

Cytokines have been classified in several different ways, although none are entirely satisfactory. *Table 4.1* show an arrangement of them grouped into six families.

Family	Examples of members[a]
Type 1 cytokines (haematopoietic)	IL-2, IL-3, IL-4, IL-5, IL-6, IL-7, GM-CSF[b]
Type 2 cytokines (interferons [IFNs])	IFN, IL-4, IL-10, IL-20, IL-22
Type 3: interleukin (IL)-1	IL-1α, IL-1β, IL-18, IL-33, IL-36
Type 4: IL-17	IL-17A, IL-17B, IL-17C, IL-17D
Type 5: tumour necrosis factor (TNF) family	TNF-α, TNF-β, BAFF[c]
Type 6: chemokines	IL-8, CCL2, CCL3, CCL19, CCL21

CCL, chemokine CC motif ligand.

[a] These are not comprehensive lists. [b] Granulocyte–macrophage colony-stimulating factor. [c] B cell survival factor.

Table 4.1 The six cytokine families. See also *Table 4.6*.

Figure 4.5 Molecular model of a B cell activating factor (BAFF). PDB file 5Y9J.

Type 1 is the largest of the six families of cytokines. All share a common protein structure consisting of a four α-helical bundle. They stimulate a wide array of immune activities, including promoting cell proliferation and differentiation, and the secretion of antibodies (*Section 4.7*).

Type 2 cytokines (interferons [IFNs]) comprise three subfamilies:

1. type I IFNs, that is, IFN-α and IFN-β
2. type II IFN or IFN-γ
3. type III IFN or IFN-λ

Interferons-α and -β (*Figure 4.6*) are produced by cells that have viral infections. They disrupt viral replication in neighbouring healthy cells by inducing those cells to produce enzymes that inhibit replication of viral nucleic acid and the production of viral proteins. Type II IFN or IFN-γ is produced by activated T cells and NKCs. It activates macrophages, in addition to stimulating cells of the adaptive immune system (*Section 4.6*). Type III IFN or IFN-λ is secreted by specific types of dendritic cells; its actions resemble type I IFNs in that it is active in antiviral defence.

Cytokines of the IL-1 family are secreted by dendritic cells, macrophages and monocytes (*Section 4.6*) as an early response to viral, bacterial or parasitic infections. Both

Margin Note 4.2
Are cytokines hormones?

Like cytokines, numerous hormones stimulate target cells by binding to cell surface receptors and stimulating increases in intracellular second messengers (*Chapter 7*). However, cytokines differ from hormones in that they are not formed in endocrine glands and operate at lower concentrations, that is, at 10^{-9} to 10^{-12} mol/L, and, while most hormones act in a fairly specific manner (adrenaline is a notable exception), those of cytokines may be complex in that they are pleiotropic, show redundancy and have a synergistic effect.

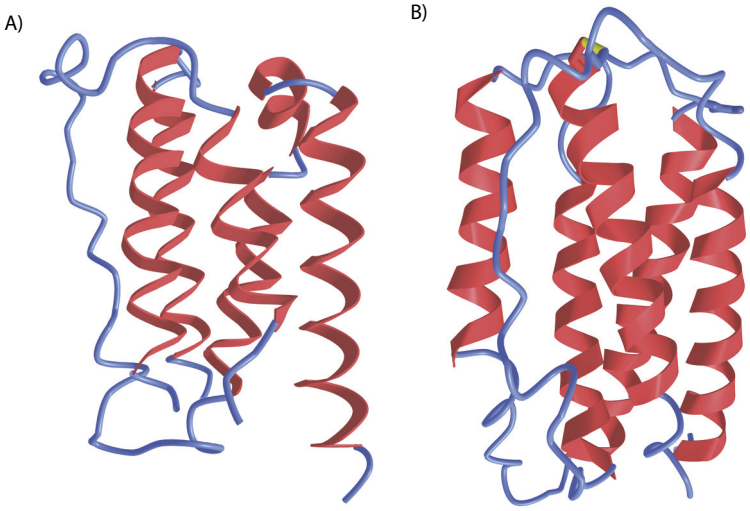


Figure 4.6 Molecular models of (A) interferon-α and (B) interferon-β. Note they both show the four helical bundle structure of cytokines. PDB files 1RH2 and 1AU1, respectively.

IL-1α and IL-1β are synthesised as proproteins (proIL-1α and proIL-1β) with M_r of 31 000. They are hydrolysed to their smaller, mature forms in reactions catalysed by the protease caspase-1, which occurs in a multiprotein complex called an **inflammasome**. Perhaps surprisingly, both proIL-1α and IL-1α are biologically active, although the former often occurs as a membrane protein. This cytokine family stimulates the liver to produce and release other cytokines, such as IFNs and chemokines (see below). Their general effects are systemic and strongly proinflammatory, leading to a general fever that helps eliminate temperature-sensitive bacteria in patients. Interleukin-1 also activates B and T cells and, therefore, is an intermediary between nonspecific and specific immune systems.

Members of the IL-17 family are also proinflammatory. For instance, IL-17A is secreted by activated T cells (*Section 4.8*) and binds to receptors on, for example, keratinocytes and neutrophils, where it induces the production and release of cytokines that promote inflammation.

Some members of the TNF family are integral membrane proteins. However, both TNF-α and TNF-β (also known as lymphotoxin-α [LT-α]) are soluble proteins. Tumour necrosis factor-α is produced by eosinophils, fibroblasts, keratinocytes, macrophages, mast cells and NKCs. It exerts multiple effects by modifying the physiological activities of a variety of cell types. Like many cytokines, it is proinflammatory and has been called an endogenous pyrogen, since, like the IL-1 family, its actions increase body temperature. Tumour necrosis factor-β/LT-α is secreted by activated lymphocytes and, like other TNFs, can affect the survival, proliferation and differentiation of target cells or, in some cases, induce apoptosis (see *Chapter 17*).

Chemokines are cytokines that stimulate **chemotaxis** in leukocytes, thus inducing their migration up a concentration gradient towards the chemokine-secreting cell/tissue. This property of chemoattraction means that the binding of some chemokines to endothelial cells lining blood vessels regulates leukocyte movements (*Section 4.6*) from blood to sites of infection in solid tissues and are therefore proinflammatory. Chemokines also help control the trafficking of B cells, monocytes and neutrophils between blood, bone marrow and peripheral tissues.

In addition to their common function of chemoattraction, chemokines also show common, homologous structural features. They consists of single polypeptides of 70–90 amino acid residues, with M_r of 8000–10 000. Their tertiary structures are restricted by four conserved cysteine (C) residues sited near their amino termini, which form intrachain disulphide bonds (–S–S–) with other cysteine residues. Chemokines are classified into four classes depending on the residues separating the two conserved cysteine residues:

1. class 1: –CXC–
2. class 2: –C–C–
3. class 3: –C–
4. class 4: –C(X_3)–C–

The X denotes a residue other than cysteine (*Figure 4.7*). Individual members of each class are referred to as ligands (Ls), for example CXCL1 and CCL5. Given the homology of their structures, it is, perhaps, not surprising that chemokines often share functional similarities. Thus, seven members of the CXC class are active in attracting neutrophils, stimulate the formation of new blood vessels (**angiogenesis**) and bind to the same receptor, CXCR2. Some ligands are able to recognise more than one type of receptor; CCL2, for instance, binds to CCR2, CCR3 and CCR5 (*Figure 4.8*).

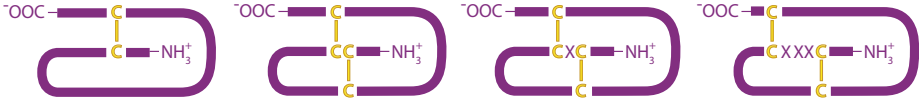


Figure 4.7 Cartoons illustrating the four classes of chemokines. C denotes cysteine and X other amino acid residues.

4.5 COMPLEMENT

Complement is a set of more than 50 plasma proteins, whose activities result in the:

- lysis of bacteria
- stimulation of the inflammatory response
- promotion of phagocytosis
- clearance of immune complexes

Margin Note 4.3
Definitely not a storm in a tea cup

Cytokine release syndrome, also called, colloquially, 'cytokine storms', arises from a dysregulation of the immune system. The definition of this syndrome is rather loose but it involves responses in which the damage done to the patient outweighs any benefits from the excessive secretion of the cytokines. Such syndromes are serious medical conditions that can cause multiorgan failure and death. For example, the massive lung damage, uncontrolled blood coagulation and many of the deaths associated with the COVID-19 pandemic are linked to the syndrome. Cytokine storms are characterised by massively increased concentrations of cytokines in the blood, particularly IFN-γ, IL-1, -6 and -18, and TNF, accompanied by a hyperactivation of immune cells. General features include a persistent severe fever, increases in ferritin concentrations and erythrocyte sedimentation rate, and thrombocytopenia (*Chapter 13*). Cytokine storms are triggered by several medical problems, including autoimmune conditions (*Chapter 5*), some cancers (*Chapter 17*), genetic predispositions or a variety of different pathogens (*Chapter 2*), and, in some cases, **iatrogenic** causes.

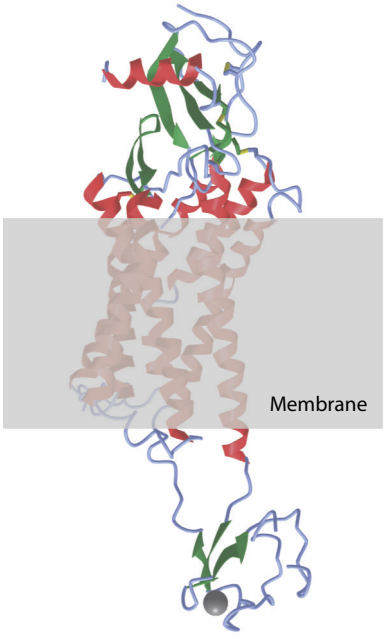


Figure 4.8 Molecular model of the chemokine receptor CCR5. PDB file 7F1T.

These proteins are often referred to as C proteins and are numbered (*Table 4.2*), for example C1q, C2 and C3, although others in the group such as mannose-binding lectin (MBL; see *Figure 5.15*) and various factors obviously do not follow this nomenclature. Complement proteins work in a sequential cascade and at certain stages individual complement proteins activate numerous molecules of the next component in the sequence. The overall effect is to greatly magnify the final immunological effect.

Protein
C1q
C1r
C1s
C2
C3
C4
C5
C6
C7
C8
C9
Mannose-binding lectin (MBL)
MBL-associated serine protease 2 (MASP-2)
Factor B (FB)
Factor D (FD)
Factor H (FH)
Factor P (FP)/properdin

Table 4.2 Examples of complement proteins.

Margin Note 4.4
Amplification is a common biochemical ploy

Amplification of initial signals or stimuli occurs often in biochemical systems, for example in the signal transduction pathways associated with cytokines and some hormones (*Chapter 7*) and in blood clotting (*Chapter 13*).

Activation of complement

Activation of complement is achieved initially in one of three ways:

1. classical pathway
2. lectin or MBL pathway
3. alternative pathway

However, all three pathways converge to a final or terminal pathway that forms a **membrane attack complex** (MAC).

The classical pathway for complement activation initially involves C1, which consists of a single C1q and two each of C1r and C1s. Initiation begins with the specific binding of immunoglobulin G (IgG) or M (IgM) antibodies (*Section 4.7*) to the cell walls of bacteria and yeast pathogens. This pathway may therefore take several days to become effective if specific antibodies are not already present. The binding of antibody to targets induces a conformational change in their molecules, allowing C1q to bind to them. C1q has several receptor sites, allowing it to bind to multiple antibodies (*Figure 4.9*). It requires at least two of these sites to bind to the antibodies on the pathogen to activate complement. For this reason, IgM, which has several sites per molecule able to interact with C1q, is a more efficient activator of complement than IgG. Indeed, a single molecule of IgM can activate complement, whereas it takes about 1000 molecules of IgG to achieve the density required for activation. The binding of C1q activates C1r and this, in turn, activates C1s, which acquires proteolytic activity. Substrates of C1s are C4 and C2, each of which is hydrolysed to two fragments: C4a and C4b, and C2a and C2b, respectively. The proteins

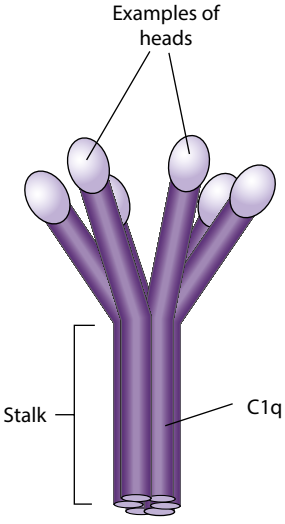


Figure 4.9 Cartoon of a C1q molecule showing the six sites ('heads'), each of which can bind to an immunoglobulin G (IgG) or M (IgM) molecule.

C4b and C2a combine to form a C4b2a complex as outlined in *Figure 4.10*. This complex is a C3 convertase, which catalyses the cleavage of C3 into C3a and C3b. This is an amplification stage in the response, given that each C3 convertase can process many substrate molecules. Protein C3 is a key component of the system and, indeed, is present in the plasma at greater concentrations than other complement proteins.

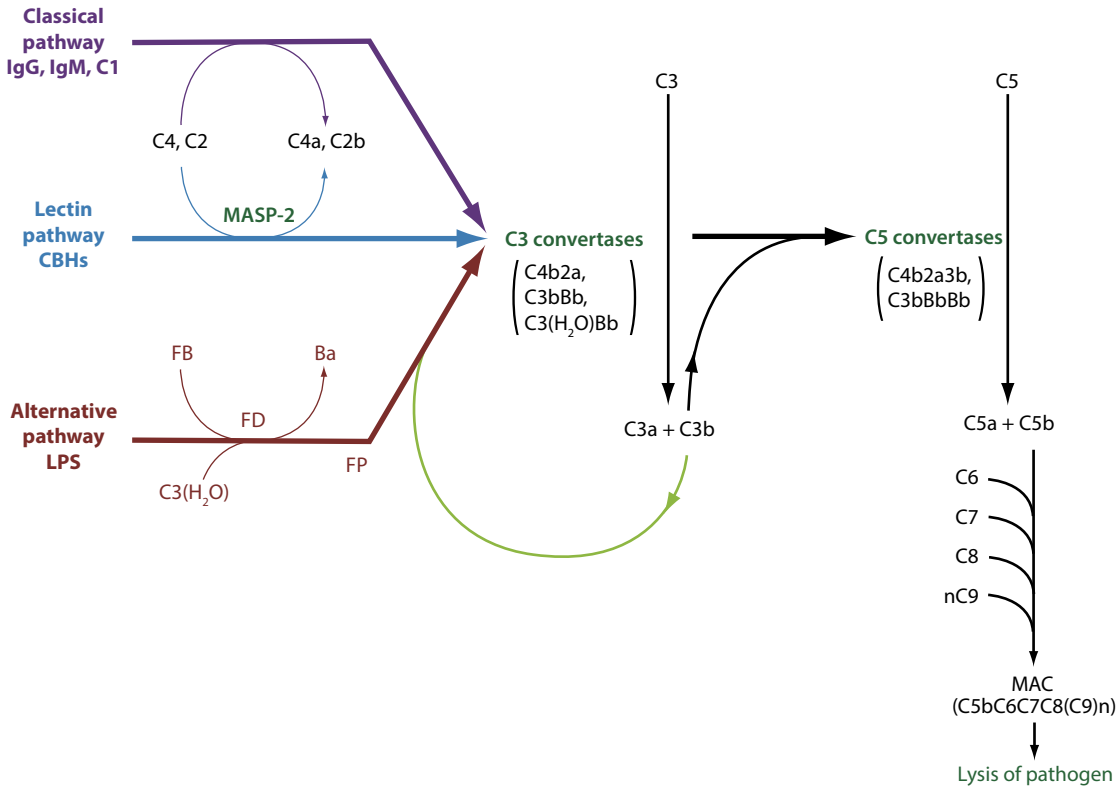


Figure 4.10 The actions of the complement pathways in forming membrane attack complexes (MACs). CBHs, carbohydrates; LPS, lipopolysaccharides.

In the lectin pathway, complement is initiated in the absence of antibodies by MBL (see *Figure 5.15*), a plasma glycoprotein produced by the liver during the early or acute stage of an infection. As its name implies, MBL is a glycoprotein that binds to mannose residues of bacterial cell walls and some viruses, although it can also recognise and bind to other sugar residues. It is thought that the hydroxyl groups on these sugars constitute a **pathogen-associated molecular pattern** (PAMP), which is recognised by the MBL. In this case, MBL therefore forms the complementary **pattern recognition receptor** (PRR). The binding of MBL activates an associated serine protease, MBL-associated serine protease-2 (MASP-2). This, in turn, catalyses the hydrolysis of C2 and C4 to form the C4bC2a/C3 convertase (see *Figure 4.10*).

The alternative pathway is also activated in the absence of antibodies by cell wall components of bacteria, such as lipopolysaccharides. Activation can be effected in three separate ways but all involve factor B (FB). In one alternative pathway, factor D (FD) catalyses the proteolysis of FB to form Ba and Bb. The latter combines with C3b to form the complex C3bBb, which is an active C3 convertase (see *Table 4.2* and *Figure 4.10*). However, the alternative pathway can also involve factor P (properdin), which binds directly to microbial membranes and recruits and stabilises the binding of the C3bBb/C3 convertase to the membrane. Alternatively, the spontaneous hydrolysis of an internal thioester bond in the C3 molecule forms small amounts of a $C3(H_2O)$ derivative in the plasma. This derivative binds FB, which is then susceptible to cleavage by FD, forming a soluble C3 convertase, $C3(H_2O)Bb$, with the release of Ba.

The final stages of all three pathways are initiated by the formation of C5 convertases, which are formed by the binding of C3b molecules to the C3 convertases: forming C4b2aC3b in the classical and lectin pathways and C3bBb3b in the alternative pathway. The final or terminal pathway leads to the production of MACs, which causes lysis of the pathogen (see *Figure 4.10*). The catalysed hydrolysis of C5 by the C5 convertases produces C5a and C5b. Again, since each enzyme can process many substrate molecules, the initial reactions are amplified. The C5b and C6 proteins bind together, exposing a C7

binding site on its surface, allowing a C5bC6C7 complex to form. This complex can insert into the surface membrane of the pathogen where it forms a receptor for C8. Binding of C8 firmly anchors the C5bC6C7C8 complex to the membrane. The C5bC6C7C8 complex, in turn, binds up to 19 C9 molecules. The C5bC6C7C8$(C9)_n$ complex is the fully formed MAC, with the $C9_n$ forming puncture pores in the membrane (*Figure 4.11*). The influx of solutes and water through these pores results in the death of the pathogen by an osmotically induced lysis.

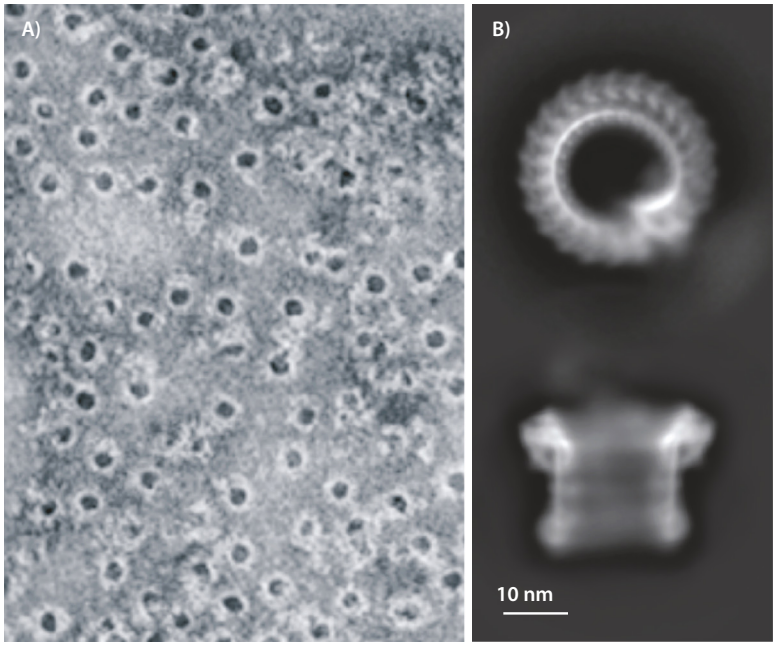


Figure 4.11 Electron micrographs showing (A) membrane attack complexes (MACs) in a membrane and (B) single MACs solubilised from the membrane with detergent. Courtesy of Professor B.P. Morgan, Systems Immunity University Research Institute, School of Medicine, Cardiff University, UK.

Complement, opsonisation and inflammation

In addition to promoting lysis of the invading microorganisms, the deposition of C3b on the surfaces of microorganisms promotes their uptake by macrophages and neutrophils that have receptors capable of binding to the complement protein. In this case, C3b is acting as an opsonin, and the process is referred to as **opsonisation** (*Figure 4.12*).

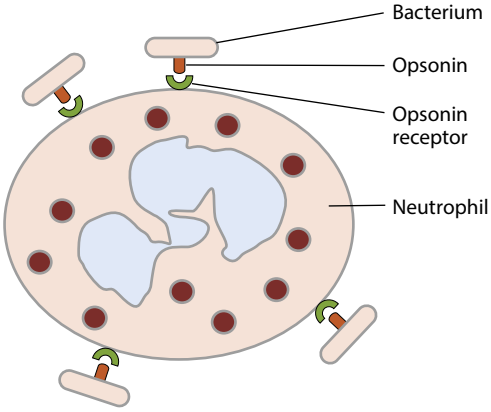


Figure 4.12 The role of an opsonin (from the Greek word *opsonin,* meaning victual [food]) in triggering phagocytosis following the binding of an opsonised bacterium to a neutrophil.

Fragments of activated complement proteins, such as C3a, C4a and C5a, called **anaphylatoxins**, promote inflammation in several ways. Their binding to receptors on a variety of cells can result in the release of the proinflammatory mediators TNF-α and IL-6, while binding to mast cells (see *Figure 5.23*) causes their degranulation with the release of inflammatory chemicals, such as histamine (*Chapter 5*). Also, anaphylatoxins increase vascular permeability and promote chemotaxis of phagocytes and mast cells to sites of inflammation (*Section 4.5*).

BOX 4.1 INFLAMMATION AND THE ACUTE PHASE RESPONSE

Inflammation describes a whole array of responses to infection and tissue damage. **Acute** inflammation is used here to describe the immediate and localised response to tissue damage. However, should an infection arise because of damage to tissues, acute inflammation can merge with the **acute phase response**. This is a whole-body response that occurs within a few hours of an infection arising. The acute phase response involves the production of a number of chemical mediators that promote and prolong local inflammation. Additionally, a **chronic** infection can lead to pronounced inflammation together with a prolonged acute phase response. This somewhat confusing series of overlapping events is (hopefully) clarified in *Figure 4.13*.

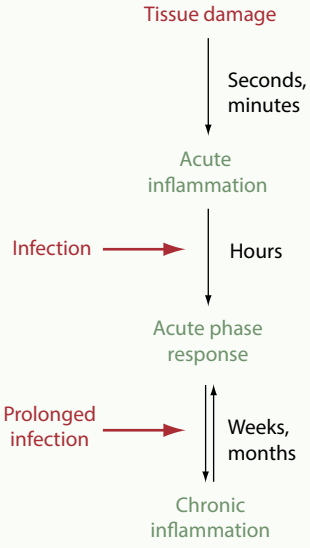


Figure 4.13 The relationship between inflammation and the acute phase response.

Acute inflammation is the rapid local response to tissue damage. For example, if the skin is scratched by a thorn, the response is a characteristic reddening and swelling of the skin, as well as sensations of heat and pain at the damaged site. This response is initiated by the release of mediators, especially histamine, from mast cells in the damaged area (*Figure 4.14*). Histamine causes blood vessels to dilate and leak plasma into the damaged tissue where it dilutes any noxious agents that entered at the damaged site and helps flush them away into the lymph. If bacteria enter the damaged area, they activate complement with the release of chemotactic proteins that encourage neutrophils to move out of the blood between the endothelial cells of capillaries.

The acute phase response occurs within hours of exposure to microorganisms. This systemic response is a primary defence against infection involving several organ systems with multiple beneficial effects. The response is brought about by cytokines, including interleukin (IL)-1, IL-6 and tumour necrosis factor (TNF)-α released by monocytes and macrophages. An acute phase response results in changes in the composition of the blood, including an increased neutrophil count and the appearance of, or increased numbers of, defence proteins called **acute phase proteins**. The appearance in the blood of a particular protein, **C-reactive protein** (CRP), which binds to bacteria and activates complement, is often used as a marker of infection. The plasma concentration of CRP prior to an acute phase response is so low as to be barely measurable; thereafter its biosynthesis increases 100- to 1000-fold.

Acute phase responses have the potential to cause harm if an infection is prolonged. For example, the liver biosynthesises acute phase proteins from amino acids released by the enzyme-catalysed proteolysis of muscle tissues. Following an acute infection, this muscle protein is rapidly replaced. However, in chronic infections, such as tuberculosis, a severe and prolonged muscle wastage (**cachexia**) may result (*Chapters 10* and 17). High body temperature is another symptom of the acute phase response. While elevated body temperatures may inhibit the replication of bacteria, fever can be dangerous, especially in children, if the temperature becomes excessive.

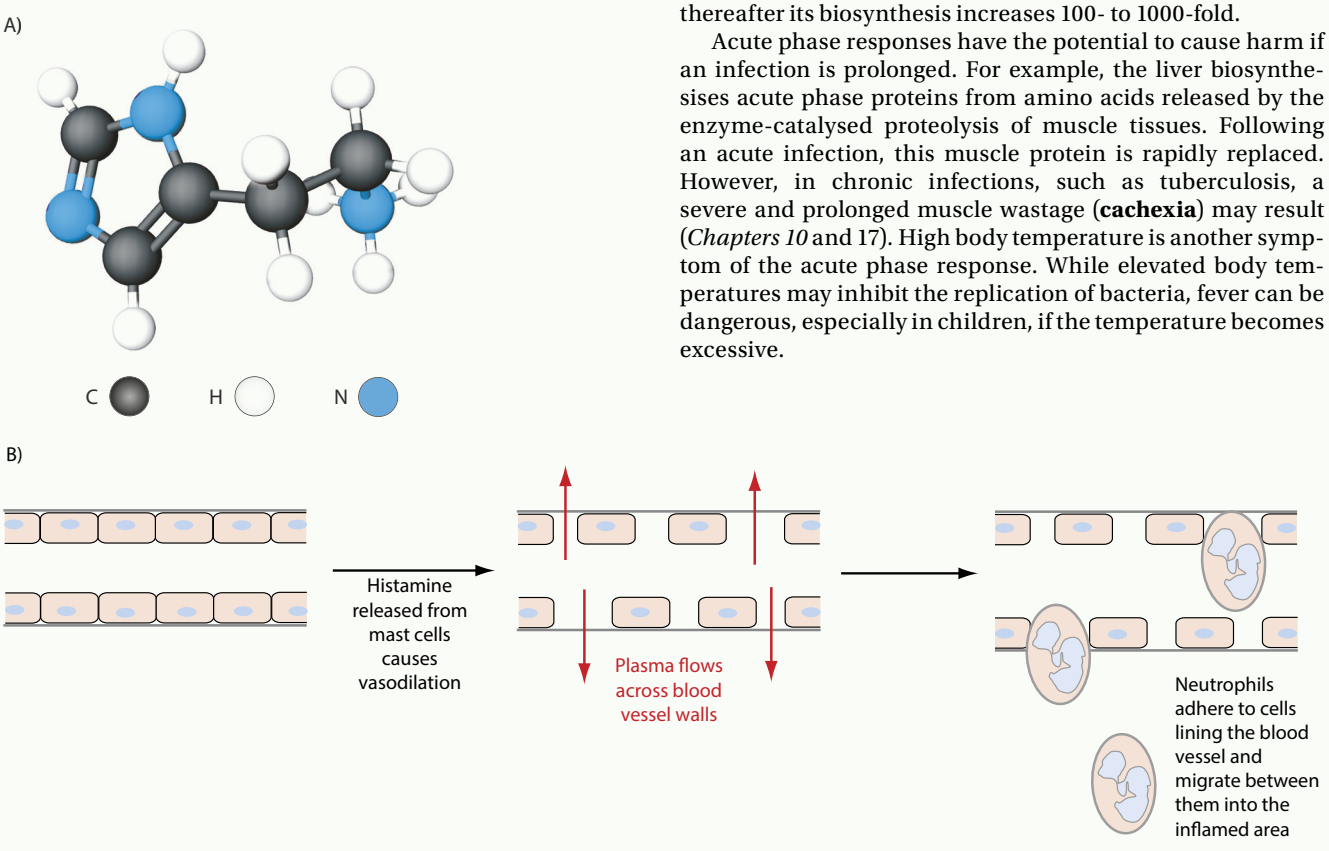

A)
C
H
N
B)
Histamine released from mast cells causes vasodilation
Plasma flows across blood vessel walls
Neutrophils adhere to cells lining the blood vessel and migrate between them into the inflamed area

Figure 4.14 (A) Computer drawn model of histamine. (B) A schematic illustration of acute inflammation.

4.6 LEUKOCYTES

Leukocytes (*Figure 4.15*) are the white blood cells produced from precursor stem cells in the bone marrow (*Chapter 13*). All have immunological roles. Approximately 80% of leukocytes are involved in nonspecific defence but the small lymphocytes are the cells of the specific immune system whose products control the numbers and activities of non-specific cells. Hence the two systems are interlinked.

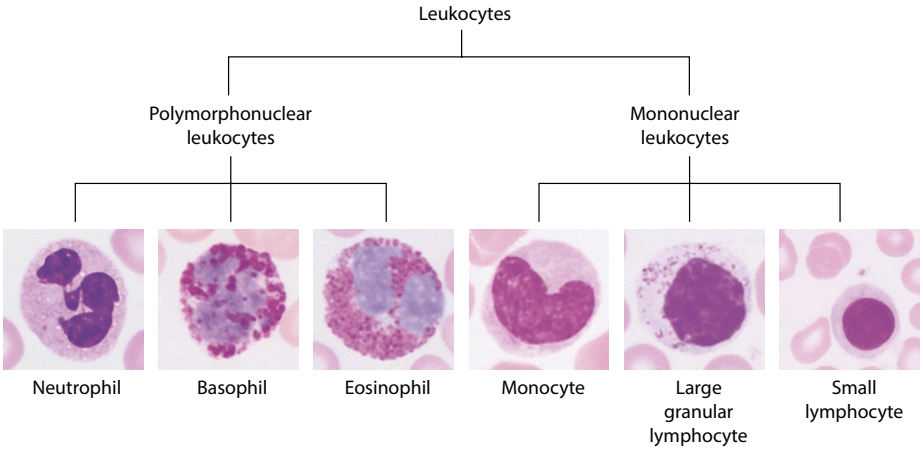


Figure 4.15 Photomicrographs of blood leukocytes. See also *Figure 13.2.* ⏎

All leukocytes can be classified into one of two groups:

1. polymorphonuclear leukocytes (PMNs)
2. mononuclear leukocytes (MNLs)

Polymorphonuclear leukocytes

Polymorphonuclear leukocytes (see *Figure 4.15*) have lobed nuclei and granular cytoplasm. They are subdivided into three groups, **neutrophils**, **basophils** and **eosinophils**, that form approximately 65% of all blood leukocytes.

Neutrophils

Neutrophils (see *Figure 4.15*) constitute approximately 60% of blood leukocytes. They are phagocytic cells that ingest and kill bacteria by opsonisation. Neutrophils have receptors for antibodies and for the activated complement protein C3b (*Section 4.5*). Thus, neutrophils can bind to bacteria coated with these proteins and phagocytosis is promoted (see *Figure 4.12*). The killing of ingested bacteria is achieved through several different mechanisms, including the use of lysosomal enzymes (*Chapter 16*); production of antibacterial chemicals, such as hydrogen peroxide, hypochlorite and nitric oxide; and use of cytoplasmic proteins called defensins (*Figure 4.16*) that attack the membranes of the ingested microbe.

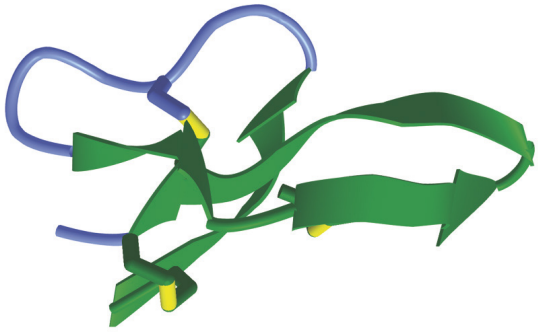


Figure 4.16 Molecular model of a defensin protein. PDB file 1ZMM. ⏎

Basophils (and mast cells)

Basophils (see *Figure 4.15*) have prominent cytoplasmic granules that take up basic stains such as toluidine blue. Basophils occur in low numbers in the blood, usually forming less than 1% of the leukocytes population. Their granules are rich in pharmacologically active agents, such as histamine and heparin, and factors that are chemotactic for other PMNs (*Table 4.3*); thus, basophils are essential for initiating inflammation despite their relatively low numbers. In addition, they have the capacity to biosynthesise and secrete other types of immune mediators when appropriately stimulated.

Chemical	Activity promoted
Histamine	Dilates blood vessels, increases vascular permeability, promotes contraction of smooth muscle cells
Heparin	Inhibits clotting of blood
Chemotactic factor for neutrophils (NCF-A)	Attracts neutrophils into inflammatory area
Chemotactic factor for eosinophils (ECF-A)	Attracts eosinophils into inflammatory area
Proteases	Degrade basement membrane of blood vessels, promoting migration of cells associated with inflammation into that area
Serotonin	Increases vascular permeability and promotes contraction of smooth muscle

Table 4.3 Some pharmacologically active chemicals in basophil granules. ⏎

Mast cells resemble basophils but, rather than occurring in the blood, are found in solid tissues throughout the body, especially in the skin, mucosal membranes and epithelia of the respiratory and gastrointestinal tracts, and in the connective tissue of a variety of internal organs. Mast cells are of greater clinical significance than basophils. Their cytoplasm is highly granular and contains similar mediators to those of basophils, although the granular contents may vary according to location. Mast cells also secrete several cytokines, many of which are proinflammatory. Both mast cells and basophils have receptors for the class of antibodies called immunoglobulin E (IgE; *Section 4.7*), which has clinical significance in the pathology of some of the most common allergic immunological disorders (*Chapter 5*).

Margin Note 4.5
Possible problems associated with basophils

The primary, that is, granular, and secondary induced mediators produced and released by basophils are essential for initiation, promotion and maintenance of inflammation. However, their inappropriate release can result in immunological disorders such as hay fever and asthma (*Chapter 5*).

Eosinophils

Eosinophils (see *Figure 4.15*), like basophils, have prominent cytoplasmic granules but, because these contain highly basic proteins, they take up acidic stains, for example eosin. They are also found in relatively low numbers in the blood, usually forming less than 2% of leukocytes. Eosinophils can be attracted to areas of inflammation by chemotactic factors secreted by basophils and mast cells. Their major role is to assist in the elimination of multicellular parasites such as tapeworms and nematodes (*Section 4.9* and *Chapters 2* and 3) by first binding to the surfaces of these parasites, often using antibodies, and then secreting their toxic granular proteins onto the surface. Eosinophils are also capable of defence.

Mononuclear leukocytes

Mononuclear leukocytes (see *Figure 4.15*) have more rounded nuclei than the PMNs. They also are subdivided into three distinct groups: **monocytes** and **large granular lymphocytes (**LGLs**)**, which are nonspecific immune cells, and **small lymphocytes**, which are responsible for the specific immune responses and are described in *Section 4.8*.

Monocytes (and macrophages)

Monocytes (see *Figure 4.15*) have characteristic indented, often horseshoe-shaped, nuclei and granular cytoplasm. They make up approximately 5% of the blood leukocytes. Monocytes are immature cells that circulate for only a matter of hours before they enter solid tissues and develop into macrophages. The spleen, lungs, liver, lymph nodes and

tonsils contain especially high numbers of macrophages. Monocytes and macrophages are phagocytic cells, clearing the blood and solid tissues of microbes and dead or dying host cells, including neutrophils and erythrocytes. Their numbers and generalised distribution ensure they effectively clear foreign material.

Macrophages have killing mechanisms resembling those of neutrophils and can be stimulated by target cells coated with antibodies and/or complement proteins. Following phagocytosis of microorganisms, especially bacteria, they secrete a range of cytokines, including IL-1, -6 and -8, and TNF-α (*Section 4.4*). Interleukin-1, IL-6 and TNF-α initiate events in the early or acute phase of infection, while the chemokine IL-8 attracts neutrophils to its source.

Monocytes and macrophages also play significant roles in the specific immune response. They are able to 'process' foreign material and 'present' it to certain types of small lymphocyte in a form they can recognise. As such, they are **antigen presenting cells** (APCs), although they are not the only cells to carry out this activity (see *Section 4.8*).

Large granular lymphocytes

Large granular lymphocytes (see *Figure 4.15*) have rounded nuclei and granular cytoplasm. They constitute 5–10% of blood leukocytes. Functionally, LGLs present a mixed population of cells: some are NKCs that kill virally infected cells and some types of tumour cells nonspecifically (*Figure 4.17*). Natural killer cells bind to the target cell and release proteins, some of which perforate the target cell membrane, while others induce a genetically programmed cell death called **apoptosis** (*Chapter 17*). Natural killer cells are the first line of defence against viruses, since they prevent their replication and spread. They may also form a defence against potential tumours by destroying some cancerous cells as they arise in the body (*Chapter 17*).

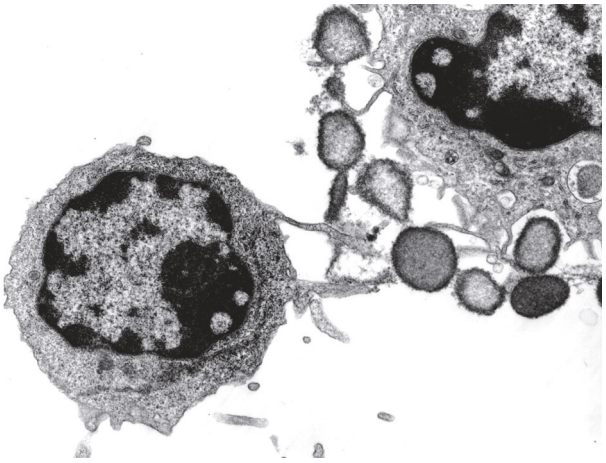


Figure 4.17 Electron micrograph of a natural killer cell (NKC; bottom left) destroying a target tumour cell. Courtesy of Dr J.T. Thornthwaite of the Cancer Research Institute of West Tennessee, USA. ⏎

Small lymphocytes

Small lymphocytes, as their name implies, are the smallest of the leukocytes (see *Figure 4.15*). They have a spherical nucleus surrounded by a paucity of cytoplasm. Small lymphocytes are the cells responsible for specific immunity. They form approximately 20% of blood leukocytes. The precursors of small lymphocytes originate in the bone marrow by division of **lymphoid stem cells** (see *Chapter 13*). However, they mature at different sites in the body to form two distinct populations that have different functions. Some remain in the bone marrow where they mature into **B lymphocytes**; others mature in the thymus into thymus-dependent or **T lymphocytes** (*Figure 4.18*).

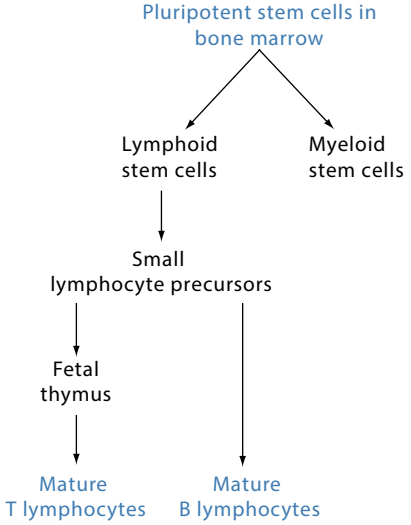


Figure 4.18 Schematic showing the development of small (B and T) lymphocytes. ⏎

4.7 SPECIFIC IMMUNE RESPONSES

Specific immune responses facilitate the development of true immunity against infectious agents. They can develop only after exposure to a virus or microorganism or a harmless vaccine derived from it; true immunity is an **acquired immunity**. For example, an individual who has had measles is unlikely to suffer that disease again, even when subsequent exposures to the virus occur.

Two major features define a specific immune response. First, specific immunity is only induced towards the **immunogen** that stimulated it. Second, the specific immune

response is more rapid following subsequent contacts with that same immunogen. It is the rapid, secondary response that prevents the development of disease in the immune individual. The ability to produce a quicker response on second contact with an immunogen is known as **immunological memory**. This process is mimicked in the use of vaccines, where, in traditional approaches, an attenuated or less virulent strain of a virus or bacterium is administered to induce immunity without causing the disease (but see also *Chapter 3*).

Immunogens may be single proteins, glycoproteins, lipoproteins, viruses or whole organisms. In general, proteins with M_r greater than 5000–10 000 are immunogenic. Polysaccharides are usually only weakly immunogenic, although they often become more immunogenic when conjugated to a protein. Whole immunogens are not recognised by cells of the specific immune system; rather, they respond to small regions of immunogenic macromolecules known as **epitopes**. Usually, epitopes are sequences of 5–7 amino acid residues in immunogenic proteins but they may also be short sequences of sugar residues in polysaccharides, lipopolysaccharides or glycoproteins called glycans. A large macromolecule may have many epitopes that are recognised by the immune system. However, a protein that has some similarities with self proteins, for example bovine serum albumin that has a similar size and structure to human albumin, will have fewer nonself regions recognisable as epitopes.

The term **antigen** is often confused with an immunogen and, indeed, some textbooks use the terms interchangeably. Here, an antigen is defined as something that will bind to the product of an immune response, such as an antibody. The term is used mostly when discussing immunotechniques, where an antibody is used to detect an antigen or quantify how much of it is present in a sample (see *Box 4.2*).

Margin Note 4.6
Haptens

Some molecules, such as steroid hormones, are too small to stimulate a specific immune response. However, if such molecules are attached covalently to an immunogenic protein, they can be recognised by the immune system and an immune response will then be mounted against them. In this case, the small molecule is called a **hapten** and the protein to which it is attached is referred to as the carrier. The ability of the immune system to recognise small molecules when they are presented in the appropriate context has proved extremely useful because it allows antibodies to be produced against a wide range of small molecules. These antibodies can then be used in specific assays to determine the presence and/or concentrations of the hapten in biological samples (see *Box 4.2*). For example, assays for steroid hormones use antibodies that have been produced in animals by immunising them with the steroid linked to a protein.

4.8 HUMORAL AND CELL-MEDIATED IMMUNITY

The specific immune response to an infectious agent, such as rubella virus (*Figure 4.19*), involves both humoral and cell-mediated immunity. Indeed, most immunogens stimulate both humoral and cell-mediated immunity. Humoral immunity involves antibodies found in body fluids, including blood and lymph, and in body secretions such as saliva, tears, mucus, colostrum and milk. It is therefore most effective at eliminating microorganisms that live extracellularly in the host. Cell-mediated immunity, on the other hand, is effective against intracellular parasites, which includes all viruses and a number of intracellular bacteria, such as *Mycobacterium tuberculosis* (see *Figure 3.12*) and *Listeria monocytogenes*. However, antibodies are useful in protecting cells in the initial stages of a viral infection and in preventing the spread of viruses from one cell to another.

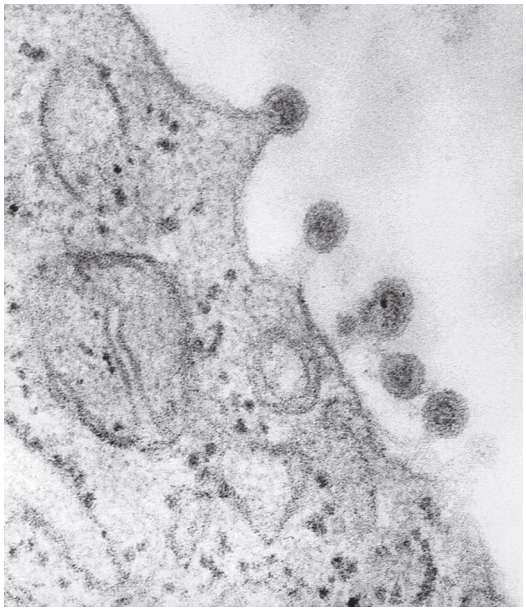


Figure 4.19 Electron micrograph of rubella virions, the causative agent of German measles, being released from an infected cell. Courtesy of the Centers for Disease Control and Prevention/ Dr F. Murphy and S. Whitfield.

Humoral immunity

A humoral response produces specific antibodies, which are glycoproteins found mostly in the γ globulin fraction of blood plasma (*Figure 4.20*). Antibodies have binding sites that are complementary to an epitope. This allows them to bind to the immunogen and initiate its destruction by, for example, complement, phagocytic cells and LGLs. The strength of this binding between the immunogen (or antigen) and antibody is referred to as **affinity**. **Avidity**, an easily confused term, is a measure of the ability of an antibody to bind to multiple binding sites.

Collectively, antibodies are known as **immunoglobulins (Igs)**. All Igs are polymeric and composed of combinations of *heavy* polypeptides, usually referred to as *chains* by immunologists, with M_r of approximately 50 000, and *light* chains of about M_r 25 000, which are joined by disulphide bonds. The chains are named using Greek letters. There are five types of Ig heavy chains called α (two subtypes), δ, Є, γ (four subtypes) and μ and two types of light chains: κ and λ. Immunoglobulins are a heterogeneous collection of molecules indeed, given that each antibody is specific for an individual epitope. Despite this heterogeneity, they can be divided into five major classes or isotypes:

1. IgG
2. IgM
3. IgA
4. IgE
5. IgD

This classification is based on differences in their structures and, particularly, on which type of heavy chain is present. *Table 4.4* lists some of their properties and differences.

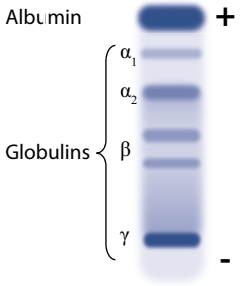


Figure 4.20 Separation of serum proteins (albumin and globulins) by electrophoresis. ⏎

Class	Subclasses	M_r	Heavy chains	Activates complement?	Crosses placenta?	Opsonisation?	Triggers inflammation?
IgM	None	900 000	μ	Yes	No	No	Indirectly via complement
IgG	IgG_1	146 000	γ_1	Yes	Yes	Yes	Indirectly via complement
	IgG_2	145 000	γ_2	Yes	Yes	Poorly	
	IgG_3	170 000	γ_3	Yes	Yes	Yes	
	IgG_4	146 000	γ_4	Restricted	Yes	Yes	
IgA	IgA_1	160 000 to 380 000	α_1	No	No	Yes	No
	IgA_2		α_2	No	No	Yes	No
IgE	None	188 000	ε	No	No	No	Directly by binding to ε receptors on mast cells
IgD	None	184 000	δ	No	No	No	No

Table 4.4 Properties of immunoglobulin (Ig) classes. Light chains are always κ or λ. ⏎

The concentrations of Igs in plasma given below represent their mean values. During infections, concentrations increase as humoral immunity is activated. Values significantly above the normal range for an appropriate age are associated with myeloma, a cancer of antibody-producing cells (see *Figure 1.9*) or infection with Epstein–Barr virus (*Chapter 3*). Concentrations significantly below the normal range are associated with immunodeficiency disorders (*Chapter 5*).

Immunoglobulin G

Immunoglobulin G is a symmetrical molecule consisting of four polypeptides of which two are identical heavy and two are identical light chains joined together by disulphide bonds, as shown in *Figure 4.21*. Both heavy and light chains have a highly conserved domain structure. These domains are globular regions of approximately 110 amino acid residues and are stabilised by intrachain disulphide bonds (*Figure 4.22* and *Margin Note 4.7*). The heavy chain of IgG has four domains while the light chain has only two.

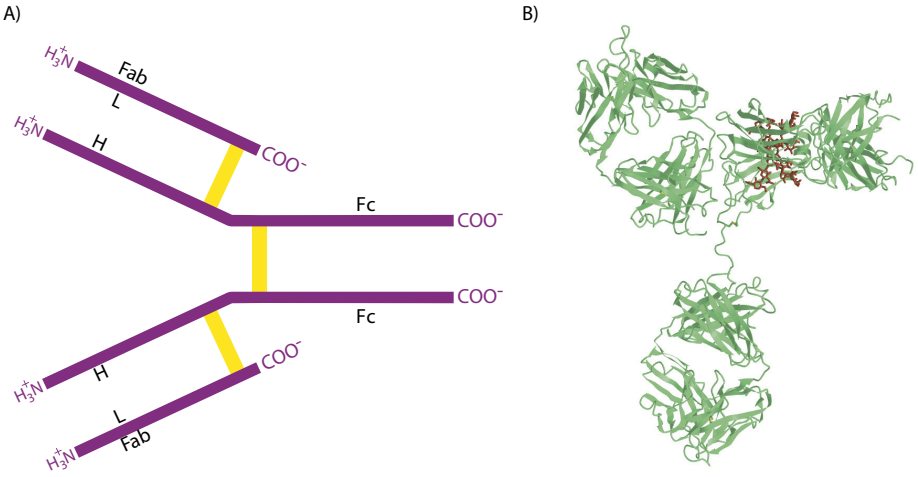


Figure 4.21 The structure of immunoglobulin G (IgG). (A) Diagram with disulphide bonds indicated in yellow. (B) Molecular model with its glycans shown in brown. PDB file 1HZH.

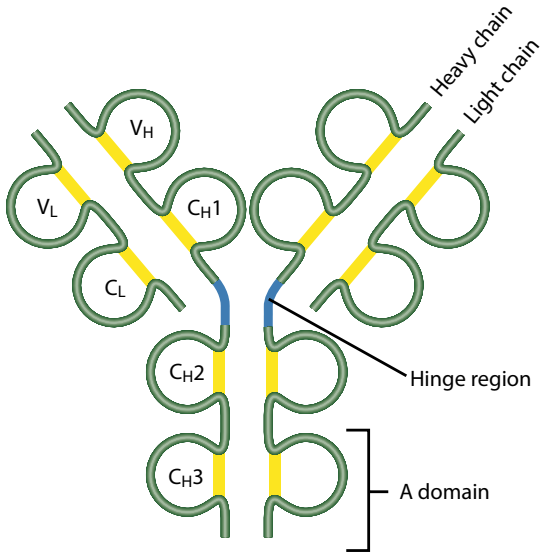


Figure 4.22 A schematic showing the domains of an immunoglobulin molecule. Disulphide bonds stabilising the homologous domains are shown in yellow; the interchain ones have been omitted for clarity.

The interchain disulphide links between the two heavy chains and between the light and heavy chains (see *Figure 4.21*) produce two units: the fragment antigen binding (Fab), of which there are two, and a single fragment crystallisable (Fc) portion. Part of the Fab can recognise and bind to an epitope (*Figure 4.24*). The Fc portion consists only of the carboxyl terminal halves of the heavy chains and is concerned with complement activation, placental transfer and binding to LGLs and phagocytes.

At concentrations of approximately 13.5 mg/mL, IgG is the most abundant antibody in blood and lymph. In humans, IgG occurs as four subclasses, all with slightly different properties and roles in the body. For example, all subclasses except IgG_4 can activate complement (see *Table 4.4*). It is evenly distributed between the vascular and extravascular compartments and is also the only antibody in humans that can cross the placenta, allowing maternal antibodies to protect the developing fetus from infection. On some occasions, this may cause problems, for example when the mother becomes sensitised to fetal antigens or has an autoimmune disease (*Chapter 5*).

Immunoglobulin G is produced after IgM on first exposure to an immunogen but is the predominant antibody produced in the secondary response (*Figure 4.25*).

Margin Note 4.7 Immunoglobulin domains

The domain substructure of IgG antibodies is found in proteins throughout the immune system. This indicates that these proteins share a common evolutionary relationship. Proteins with this domain structure are members of the immunoglobulin superfamily (*Figure 4.23*), which, in addition to Igs, includes T cell receptors, major histocompatibility complex (MHC) molecules and cell surface cluster of differentiation (CD) proteins.

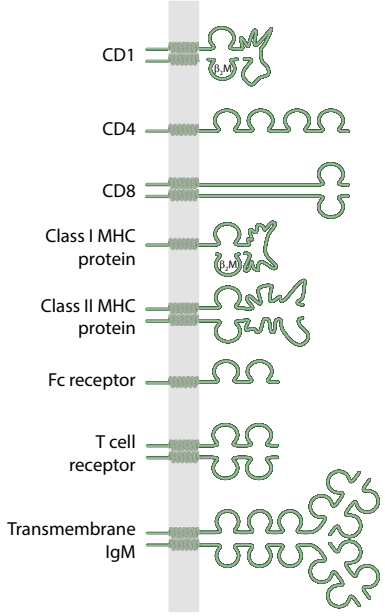


Figure 4.23 Schematic representations of members of the immunoglobulin superfamily, several of which are discussed in this chapter. The homologous domains common to all are represented as Ω shapes.

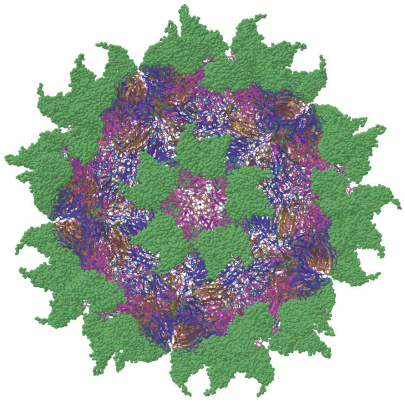


Figure 4.24 Molecular model showing the binding of immunoglobulin G (IgG) fragment antigen binding (Fab) fragments (green) to epitopes on the surface of the capsid of a rhinovirus (*Chapters 2* and *3*). PDP file 1RVF.

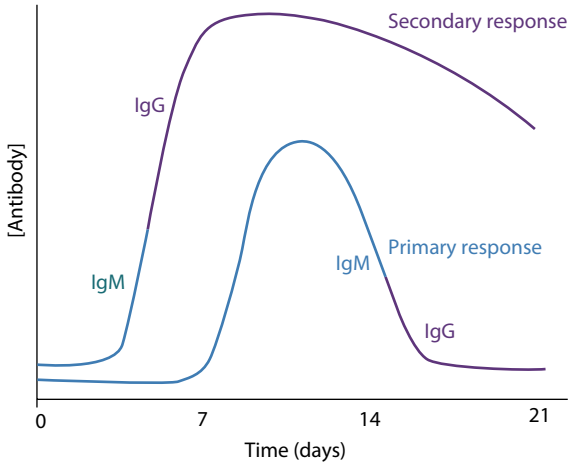


Figure 4.25 The primary and secondary antibody responses to an immunogen. Note how the first antibody produced in the primary response is immunoglobulin M (IgM), which is then superseded by immunoglobulin G (IgG). In the secondary response, IgG is the major immunoglobulin produced.

Immunoglobulin M

Immunoglobulin M is the largest of the antibodies with M_r of about 900 000. Four polypeptides form a structure somewhat resembling that of IgG, which is repeated five times. The five units are joined at their Fc portions by a protein called the J or joining chain (*Figure 4.26*).

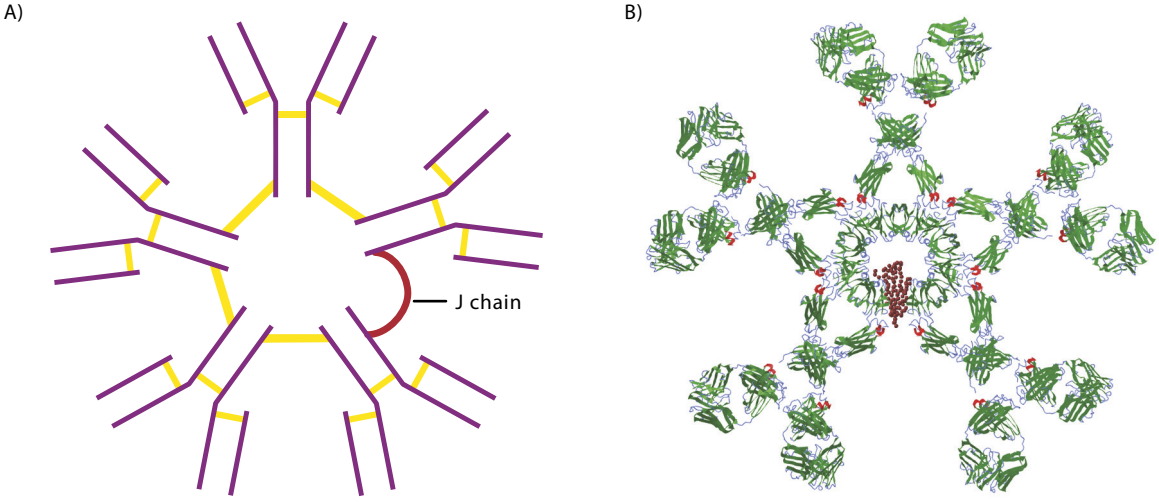


Figure 4.26 (A) Schematic illustrating the structure of an immunoglobulin M (IgM) molecule with a J holding the five units together. (B) Molecular mode of an IgM molecule, with the J chain highlighted in brown. PDB file 2RCJ.

Thus, each molecule of IgM has 10 binding sites, the largest number of any of the antibody classes, although the number in practice depends upon the size of the immunogen. This feature makes IgM efficient at agglutinating cells (clumping them together). It is also an effective activator of complement. Most IgM is vascular, with little present in lymph or secretions, constituting approximately 10% of plasma Igs at a concentration of roughly 1.5 mg/mL. Immunoglobulin M is always the first antibody to be produced during an immune response and is also the predominant one formed during a primary response (see *Figure 4.25*).

Immunoglobulin A

Most plasma IgA in humans occurs in the familiar four-chain structure akin to that of IgG with M_r of 160 000. Immunoglobulin A occurs in plasma at concentrations of 0.5–3.0 mg/mL,

a proportion of which is found in a dimeric form stabilised by J chains. The dimeric IgA is protected from enzymic attack by an additional protein known as the **secretory component** (*Figure 4.27*).

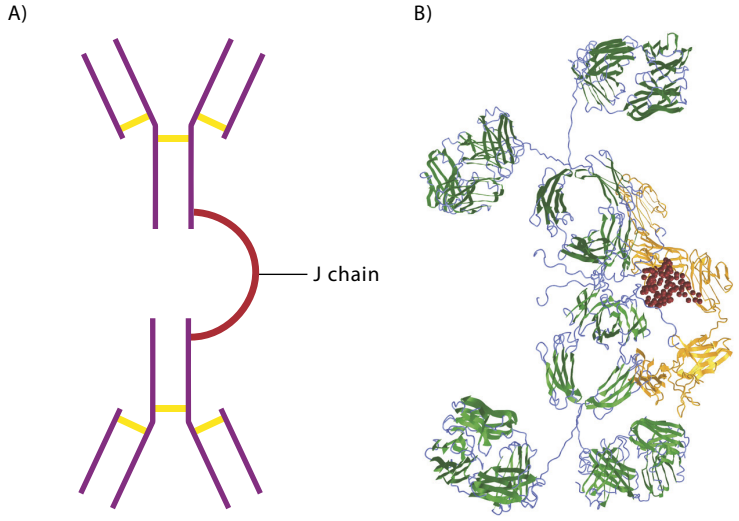


Figure 4.27 (A) Schematic showing the structure of the dimeric immunoglobulin A (IgA) molecule. The J chain holds the dimer together. (B) Molecular model of an IgA molecule with the J chain highlighted in brown and the secretory component in orange. PDB file 3CHN.

Immunoglobulin A is the major antibody found in body secretions, such as mucus, saliva and tears, and therefore protects mucosal surfaces.

Immunoglobulin E

Most IgE is found bound to the surfaces of blood basophils and tissue mast cells, since these cells possess high affinity receptors for its heavy chain. Binding to these cells prolongs the half-life of IgE from two days to several weeks. Immunoglobulin E is found in plasma at a concentration of about 5×10^{-5} mg/mL. Despite its low plasma concentrations, IgE is a potent stimulator of inflammation, since the binding of epitopes to mast cell-bound antibody can trigger degranulation of these cells. This aids the protection of external surfaces to pathogens that have not been neutralised by surface IgA. However, the proinflammatory nature of IgE is, perhaps, best seen in the elimination of multicellular parasites, such as tapeworms and nematodes that have been coated with opsins. However, in susceptible individuals, IgE can also trigger inflammatory responses to otherwise harmless immunogens, such as pollen, resulting in allergic reactions (*Chapter 5*).

Immunoglobulin D

The role of secreted IgD (see *Figure 5.8*) is uncertain. However, it does have a role in the recognition of epitopes by specific types of B lymphocytes of the immune system. Together with IgM, it may form an immunogen receptor; binding of an epitope then activates the cell. Immunoglobulin D is found in plasma at a concentration of about 30 μg/mL, although in hyperimmunoglobulin D syndrome (HIDS) these levels are increased. The syndrome is associated with periodic fever and joint disease (see *Box 5.3*).

BOX 4.2 IMMUNOASSAYS

Antibodies are powerful reagents for detecting and quantifying antigens. The specificity of antibodies allows biomedical scientists to measure the concentration of an analyte, for example a steroid hormone, in biological fluids such as plasma, which contain hundreds of other biomolecules, some of which may be similar in structure to the analyte being measured. The use of antibodies to quantify antigens is called **immunoassay**, and immunoassays are used in all branches of the biomedical sciences. Some immunoassays are among the most sensitive analytical techniques known, detecting antigens in the range of pg/mL to mg/mL.

BOX 4.2 IMMUNOASSAYS (CONTINUED)

One of the earliest immunoassays developed was *radial immunodiffusion* (RID). This method relies on the ability of antibodies to precipitate soluble protein antigens. Antibody is incorporated evenly into an agar gel and a measured volume of antigen solution is added to wells cut into the agar. As the antigen diffuses into the agar, the reaction with antibody forms a circle of precipitation (*Figure 4.28*). After allowing all of the antigen to diffuse, which takes about 72 hours, the diameter of this precipitin ring is measured. The antigen concentration is proportional to the square of the diameter so that the concentration of the unknown can be determined from a standard curve produced using known antigen concentrations.

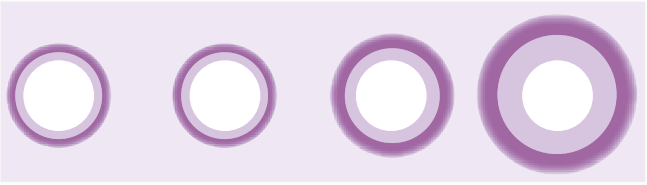


Figure 4.28 An outline of a radial immunodiffusion (RID) assay. Note that the diameters of the precipitin rings increase with larger concentrations of antigen.

Radial immunodiffusion is a simple and reliable method that can be used, for example, to measure the concentrations of a number of serum proteins. However, it is insufficiently sensitive to determine concentrations of µg/mL to mg/mL and takes several days before the result can be noted.

Nephelometry also detects precipitations, but in solution rather than in agar. It relies on the ability of precipitates formed by an antibody reacting with a protein antigen to scatter a beam of light passed through it. Scattered light is detected at right angles to the original light source. This method is much more rapid than RID, although optically clear antibodies must be used.

Methods that use 'labelled' reagents are much more sensitive and can be used to measure nonprotein antigens. In addition, results from these assays can be obtained within hours rather than days. The first of these labelled reagent techniques to be developed was *radioimmunoassay* (RIA), which was devised in 1960. Radioimmunoassay relies on the competition between radiolabelled and unlabelled antigen for a limited amount of antibody. The more unlabelled antigen (standard or unknown) there is in a sample, the less radioactive antigen will bind to the antibody. Radioimmunoassays are extremely sensitive, measuring routinely pg/mL to µg/mL. However, since they employ radioactive substances, special precautions and licensing are required and, unfortunately, the waste disposal costs are high.

Enzyme immunoassays (EIA) rely on the use of an enzyme-labelled antibody to measure an antigen. The enzyme used is one that will convert a colourless substrate into a coloured soluble product that can be measured spectrophotometrically. The most frequently used enzyme labels are horseradish peroxidase and alkaline phosphatase, and the most common EIA is the enzyme-linked immunosorbent assay (ELISA). The simplest ELISA format is to allow a protein antigen to adsorb onto the wells of a plastic microtitre plate (*Figure 4.29*). An enzyme-labelled antibody is then added, which binds to the antigen in the wells. Wells containing a large amount of antigen will contain more antibody and therefore more enzyme. The substrate for the enzyme is then added and, after a limited period, the reactions are stopped and the absorbances measured. There are many different adaptations of ELISA, which allow the measurement of nonprotein antigens or allow the sensitivity to be increased to that approaching RIA. In addition to enzymes, fluorescent and bio- and chemiluminescent labels can be used in immunoassays.

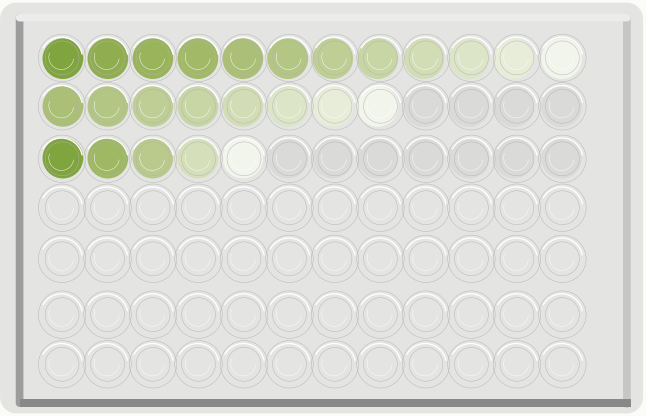


Figure 4.29 The outcome of a typical enzyme-linked immunosorbent assay (ELISA) using three sources of antigen that have each been serially diluted across the rows of wells.

Immunoassays are flexible methods, which has led to the development of several manual and automated techniques. For example, a multiplex immunoassay is a derivative of an ELISA that uses beads to bind the capture antibody, allowing the simultaneous measurements of multiple analytes in a single experiment. Multiplex assays are still more common in research facilities than clinical settings. However, many immunoassays are available for clinical purposes, for example to measure the concentrations of hormones, tumour markers, autoantibodies and allergens and to help diagnose many infectious diseases as in the lateral flow immunoassay (*Figure 4.30*) used to diagnose COVID-19 infection.

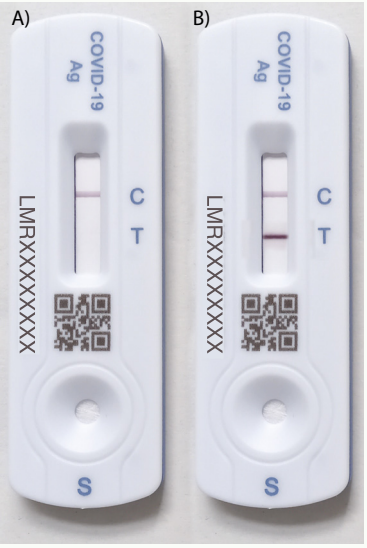


Figure 4.30 The diagnosis of COVID-19 using a lateral flow immunoassay. (A) Shows a negative result, while (B) is positive indicating the individual is infected with the SARS-CoV-2 virus.

B and T Lymphocytes

Following their production from lymphoid stem cells in the bone marrow, the precursors of mature B and T lymphocytes have different maturation paths (see *Figure 4.18*). Some small lymphocytes remain in the bone marrow where they mature into B lymphocytes. These small lymphocytes have antibodies on their surface membranes that are receptors for an individual epitope. An individual B lymphocyte is capable of clonal division and of biosynthesising and secreting antibodies specific for that epitope when it is stimulated appropriately by the immunogen. Thus, B cells are responsible for humoral immunity.

T Lymphocytes leave the bone marrow when immature and complete their maturation in the fetal thymus. This is a bilobed organ situated just above the heart. During their development in the thymus, T lymphocytes produce a cell surface receptor and therefore acquire specificity for a specific epitope. T Lymphocytes then mature into one of two major T cell subsets:

1. **cytotoxic precursor** (T_C) cells
2. **helper T lymphocyte** (T_H) cells

When appropriately stimulated, a T_C cell will develop and grow to a clone of **cytotoxic T lymphocytes** (CTLs), which can kill virus-infected cells. When stimulated by an immunogen, T_H cells develop into cytokine-secreting T_H cells that produce an array of cytokines that control the activities of both specific and nonspecific cells of the immune system. Thus, T_H cells have a central role in the regulation of all immune responses.

When mature, both B and T lymphocytes are released into the circulation. However, small lymphocytes are not confined to the blood and constantly move between the blood and the lymphoid systems (the spleen, lymph nodes, tonsils and the mucosa-associated lymphoid tissues (MALTs) found in the respiratory, gastrointestinal and urogenital tracts) in a phenomenon known as recirculation (*Figure 4.31*). Lymph, the fluid that drains from the tissues, is delivered to the blood by the thoracic duct at its junction with the left subclavian vein. En route to the thoracic duct, lymph is filtered through many lymph nodes. Small lymphocytes circulating in the blood can move between the endothelial cells lining the blood vessels that supply the lymph nodes. These vessels have a specialised endothelium that aids this process. Small lymphocytes enter the lymph node by crossing the blood vessel walls and from there they enter the lymphatic system and, eventually, return to the blood.

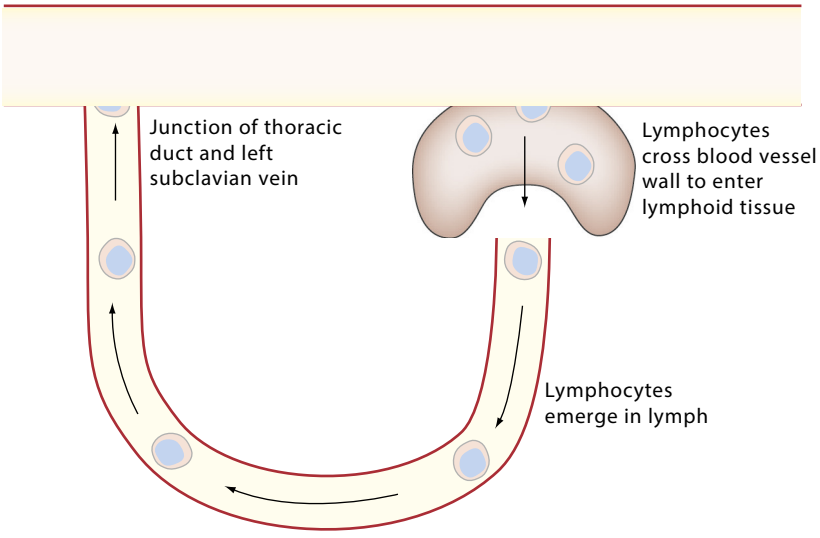

Junction of thoracic duct and left subclavian vein
Lymphocytes cross blood vessel wall to enter lymphoid tissue
Lymphocytes emerge in lymph

Figure 4.31 Schematic showing the recirculation of small lymphocytes between the blood and lymph systems.

Production of specific immune responses

Given that all small lymphocytes are specific for a single epitope on an individual immunogen, they will respond only to that epitope. The basis of this specificity depends on lymphocyte membrane proteins that act as receptors for individual epitopes. The receptors on a B lymphocyte are surface Igs that have the same specificity as the antibody

that the appropriately stimulated cell will subsequently secrete. Note that the surface antibodies on a single B cell may belong to more than one class of Ig. For example, B lymphocytes may express both IgG and monomeric IgM, but these two will have identical specificities. The T cell receptor (TCR) is composed of two polypeptides that together form a single binding site (see *Figures 4.23* and *4.34*). All the TCRs on a single T cell have the same specificity.

Although B and T cells have different types of receptor and are involved in different aspects of the immune response, there are some similarities in the ways they are activated and how they respond to a specific epitope. For example, when both B and T lymphocytes are exposed to an appropriate epitope, they enter a series of cell divisions that results in the production of a clone of cells of identical specificity (*Figure 4.32*). These cell divisions require cytokines (*Section 4.4*) that are produced and secreted by T_H lymphocytes once they have received the appropriate stimulus. Most of the cells in the small lymphocyte clone then differentiate into effector cells, the nature of which depends on the type of small lymphocyte originally stimulated. Not all the cells in the clone differentiate at this stage. Some remain as **memory cells**, awaiting the next exposure to the same epitope/immunogen when a faster and quantitatively greater response is produced.

Margin Note 4.8
Immunity and self damage

Both humoral and cell-mediated immunities protect the body from infection but, sometimes, these responses can lead to tissue damage, for example when IgE causes allergic reactions (*Chapter 5*) or when both forms of immunity help to bring about the destruction and rejection of an organ transplant (*Chapter 6*).

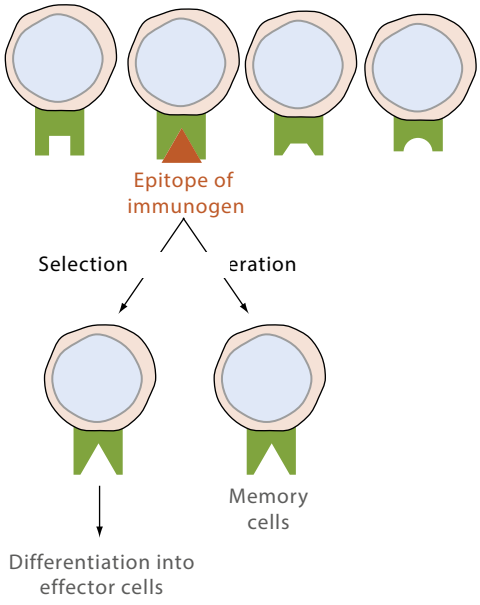


Figure 4.32 Schematic showing clonal selection in lymphocyte activation.

BOX 4.3 DISTINGUISHING B AND T LYMPHOCYTES

In a standard stained blood smear, all small lymphocytes appear similar and it is not possible to distinguish between B and T lymphocytes or between T_C and T_H cells. However, these cells can be distinguished by staining for marker proteins in the membranes of these cells. Examples of different markers that can be used to distinguish B and T lymphocyte types are shown in *Table 4.5*. The marker molecules can be distinguished by **immunohistochemistry**, in which a labelled antibody is used to stain the cells. One of the most common labels used is a fluorescent molecule, such as fluorescein, which produces an apple-green-coloured fluorescence when irradiated with light of short wavelengths.

Fluorescent-labelled antibody techniques, known as **immunofluorescence**, have been in use since the 1950s (*Figure 4.33*). They are easy to perform and reliable, although the preparations are not permanent, as the fluorescent label tends to bleach and they need to be examined soon after the test has been performed. In addition, they require the use of fluorescence microscopes, although lymphocyte samples can also be quantified by flow cytometry (see *Box 6.1*).

Protein	Role *in vivo*	Found on
Surface immunoglobulin	Epitope receptor	B lymphocytes
CD3 (see *Figure 4.34*)	Signal transduction following binding of T cell receptor to epitope	All mature T lymphocytes
CD4	Coreceptor molecule	Helper T lymphocytes (T_H)
CD8	Coreceptor molecule	Cytotoxic T cell precursors (T_C) and cytotoxic T lymphocytes (CTL)

Table 4.5 Marker proteins for B and T lymphocytes.

BOX 4.3 DISTINGUISHING B AND T LYMPHOCYTES (CONTINUED)

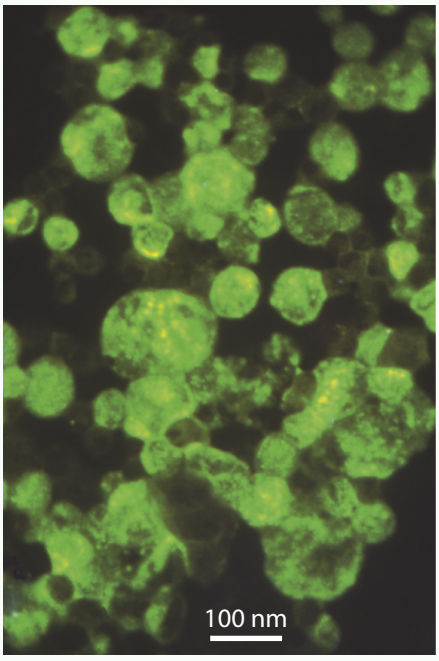


Figure 4.33 Photomicrograph of human herpes virus 2 (HHV-2; see *Table 2.5*) visualised by immunofluorescence microscopy showing the envelope fluorescing green. Courtesy of the Centers for Disease Control and Prevention /Dr C. Lyerla.

B Lymphocytes can be stained with an anti-immunoglobulin carrying a conjugated fluorescein label. This binds to the cell surface immunoglobulin, so that the B lymphocyte fluoresces when examined with a fluorescence microscope or flow cytometry.

Mature T lymphocytes can be identified by the CD3 protein carried on their surface (*Figure 4.34*). Staining usually involves an indirect method that is a two-stage process. This involves incubating the lymphocytes with an unconjugated anti-CD3 antibody, which is a monoclonal antibody originating in mice (see *Box 4.4*). This is then followed by the second stage involving incubation with a fluorescein-conjugated antibody to the mouse immunoglobulin bound to the CD3 proteins on the T lymphocyte surface (*Figure 4.34*). The cells can then be distinguished by their fluorescence. Similarly, T_H and T_C cells can be stained and identified using anti-CD4 and anti-CD8 antibodies, respectively. Labels other than fluorochromes are available, which allow the production of permanent preparations and the use of ordinary light microscopes. For example, antibodies labelled with an enzyme, such as horseradish peroxidase or alkaline phosphatase, can be located by their ability to convert a colourless substrate into an insoluble coloured compound that can be seen when the cells are examined microscopically. This technique is called **enzyme immunohistochemistry** (*Figure 4.35*).

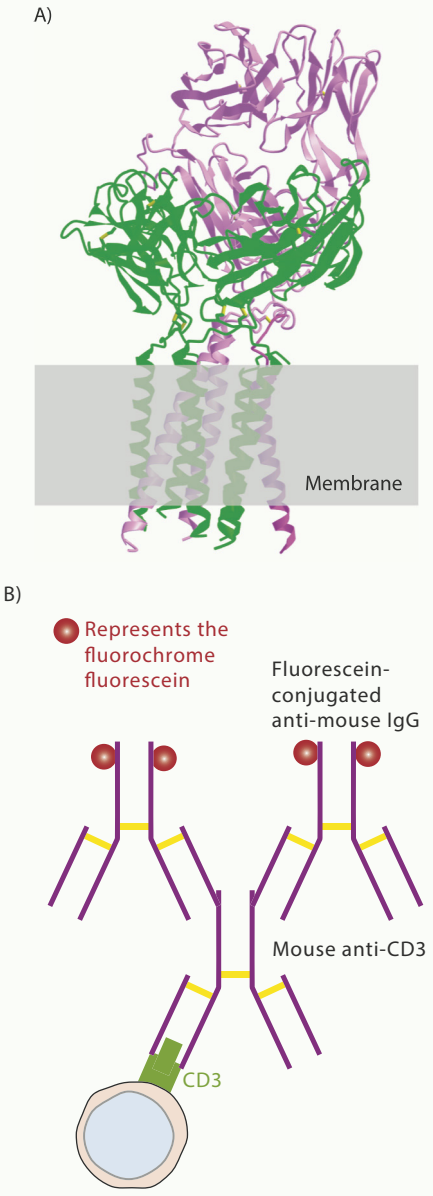


Figure 4.34 (A) Molecular model showing the membrane-bound complex of CD3 (green) and a T cell receptor (magenta). PDB file 6JXR. (B) Schematic showing the detection of cellular antigens, in this case CD3, by immunofluorescence. IgG, immunoglobulin G.

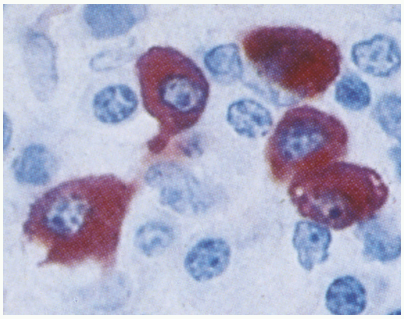


Figure 4.35 Photomicrograph showing the detection of cytoplasmic antigens following the production of a coloured marker using enzyme immunohistochemistry.

Activation of B lymphocytes in humoral immunity

The response of B lymphocytes to an immunogen is outlined in *Figure 4.36*. Once the receptors on a B lymphocyte bind to an epitope on the immunogen, the lymphocyte is stimulated to divide and differentiate under the influence of cytokines released from T_H lymphocytes, forming a clone of antibody-secreting **plasma cells** (*Figure 4.37*). Plasma cells are not found in the blood of healthy individuals. Instead, they are present in the lymph nodes, spleen and MALTs. Here, they secrete antibodies until they die after a lifespan of a few days to several months. Antibodies secreted by plasma cells in lymph nodes first appear in the lymph and then the blood, while antibodies produced in the spleen move directly into the blood. The class of antibody secreted depends, in part, on which B lymphocyte was stimulated and on the cytokines that influenced its differentiation. Certain cytokines are known to favour the production of antibodies of a particular class. For example, IL-4 promotes the production of IgE and favours a response to multicellular parasites (*Section 4.9* and *Chapter 2*). However, a predisposition to produce IgE may also make that individual more susceptible to allergic reactions, such as hay fever and allergic asthma, as will be discussed in *Chapter 5*.

Figure 4.36 shows only one B lymphocyte responding to a single epitope on an immunogen, giving rise to a single clone of plasma cells. In reality, an immunogen, such as a bacterium, possesses numerous proteins, each of which may have hundreds of epitopes. Thus, a humoral immune response involves the stimulation of many B lymphocytes, each of which can proliferate into a clone. Although each clone produces only a single type of antibody, called a **monoclonal antibody** (*Box 4.4*), hundreds of clones are formed so that the response of the system is **polyclonal**, resulting in a heterogeneous array of antibodies appearing in the blood.

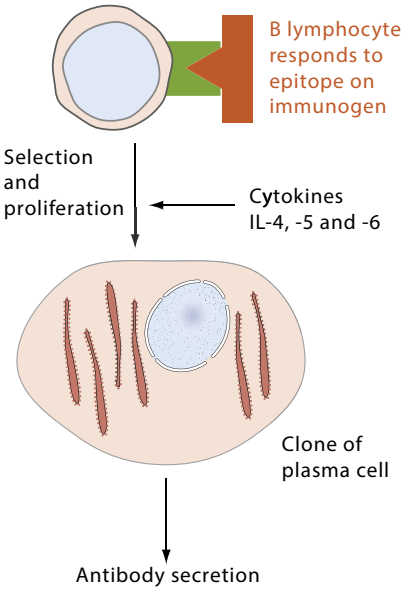

B lymphocyte responds to epitope on immunogen
Selection and proliferation
Cytokines IL-4, -5 and -6
Clone of plasma cell
Antibody secretion

Figure 4.36 Schematic illustrating the activation of a B lymphocyte and its differentiation into a plasma cell. Note that the plasma cell has more cytoplasm, with extensive rough endoplasmic reticulum synthesising immunoglobulins. IL, interleukin.

Figure 4.37 Electron micrograph of a plasma cell showing its extensive rough endoplasmic reticulum.

BOX 4.4 MONOCLONAL ANTIBODIES

Each clone of plasma cells produces homogeneous or monoclonal antibodies (MAbs) that are specific for a single epitope. In 1975, Georges Köhler (1946–1995) and César Milstein (1927–2002), working in Cambridge, UK, developed a technique whereby plasma cells producing a specified MAb could be immortalised and so grown in culture indefinitely. The technique, outlined in *Figure 4.38*, involves immunising mice with the immunogen in question. After the required immunisation protocol, the mouse spleen containing antibody-secreting plasma cells is removed and gently homogenised to form a suspension of single cells. The plasma cells are then fused with cultured cells derived from a mouse myeloma, that is, a plasma cell tumour, the cells of which are immortal. The fusing agent, or fusogen, is polyethylene glycol (PEG). The resulting cells are known as hybrid myelomas or, more frequently, **hybridomas** and are, like their myeloma progenitor, essentially immortal when grown in suspension. Since the fusogen is indiscriminate in its actions, the cell suspension also contains hybrids of plasma cell–plasma cell and myeloma cell–myeloma cell fusions. While the plasma cell–plasma cell fusions die within a short time in culture, the myeloma–myeloma fusions and any unfused myeloma cells must be selectively removed, as they would quickly outgrow the hybridomas. The most common selection procedure involves using a myeloma cell lacking hypoxanthine guanine phosphoribosyl transferase (HGPRT) activity and growing the resulting hybridomas in medium containing hypoxanthine, aminopterin and thymidine (HAT medium) for a period after hybridisation. Aminopterin inhibits dihydrofolate reductase (DHFR), which is essential for the synthesis of DNA. Cells that possess HGPRT can overcome this block by using HGPRT and thymidine kinase (TK), as long as they are supplied with hypoxanthine and thymidine. Thus, hybridoma cells, which have HGPRT supplied by the plasma cells, and TK activities survive in HAT medium, but myeloma cells do not.

A suspension of fused cells produced from a single mouse spleen may contain millions of hybridoma cells derived from

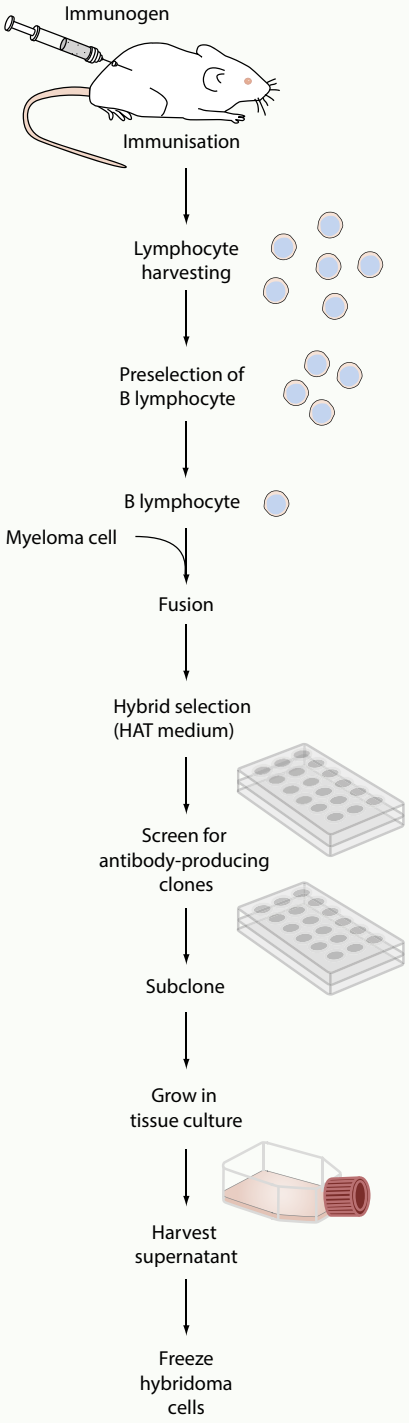


Figure 4.38 Schematic showing the production of monoclonal antibodies from the immunisation of a mouse to the culture and storage of the hybridoma cells for future use. HAT medium, medium containing hypoxanthine, aminopterin and thymidine.

different clones. To produce MAbs, individual hybridoma cells need to be isolated and grown individually. Isolation is achieved by diluting the cell suspension to such a degree that there is a high degree of certainty that aliquots will contain only a single hybridoma cell. Such aliquots can then be grown in culture to produce a clone, which will secrete a MAb.

Although hybridoma technology is still commonly used to produce antibodies, murine-derived MAbs have limited therapeutic efficacy and patients could develop a human anti-mouse antibody (HAMA) response. Antibody engineering techniques have been subsequently utilised to create chimeric or humanised antibodies by utilising the murine variable regions or complementary determining regions (CDRs), respectively, in conjunction with human constant regions, to maintain target specificity and reduce the HAMA response. Fully human antibodies are now generated using transgenic mice models, such as HuMabMouse and XenoMouse, in which the mouse immunoglobulin gene loci have been replaced with human loci within the transgenic mouse genome.

The development of antibody phage display libraries was a major technological advance, which made the discovery and development of MAbs quicker and more efficient. Phage display technology was first described in 1985. Filamentous phages were shown to express a peptide of interest on their surface after the insertion of a fragment of foreign DNA into the relevant phage coat protein gene. In 1990, phage display technology was first utilised to create a phage-displayed antibody library that could be used to discover antigen-specific MAbs. The first fully human therapeutic antibody, adalimumab, which is used to treat rheumatoid arthritis, was generated based on this technique. Its name is of interest. Man-made MAbs from mouse proteins (murine IgG) have names ending in -omab; those derived from primate proteins end in -imab. Chimeric antibodies with murine variable regions and the remainder of human origin have names ending in -ximab. Humanised antibodies that include a murine hypervariable segment only, with the rest of human origin, end in -zumab. Fully human MAbs end in -umab.

Over the last three decades, MAbs have made a striking transformation from scientific tools to powerful human therapeutics, with widespread uses in diagnosing, preventing and treating diseases. They are used in immunoassays to measure the concentrations of biomolecules in clinical samples (see *Box 4.2*). For example, commercial MAbs to hormones such as thyroxine, oestrogen and testosterone can confirm suspected hormonal deficiencies (*Chapter 7*). Monoclonal antibodies for the oestrogen and progesterone receptors and human epidermal growth factor receptor 2 (HER2) are used to diagnose subtypes of breast cancers allowing for different tailored downstream treatments (*Chapter 17*). In 2021, the United States Food and Drug Administration (FDA) authorised evusheld (tixagevimab co-packaged with cilgavimab and administered together) for the pre-exposure prophylaxis of COVID-19. The use is restricted to adults and paediatric individuals over 12 years of age and weighing at least 40 kilograms who have moderately to severely compromised immune systems due to a medical condition or immunosuppressive treatments (*Chapter 6*), and thus may not be able to mount an adequate immune response to COVID-19 vaccines.

Therapeutic antibodies have become the predominant class of new drugs developed in recent years for the treatment of human diseases. For example, adalimumab is used to treat rheumatoid and psoriatic arthritis, Crohn disease and uveitis. Natalizumab is an anti-α4 integrin MAb used for treating relapsed multiple sclerosis. Pembrolizumab, a humanised monoclonal immunoglobulin G4 antibody directed against human programmed cell death protein 1 (PD-1) on lymphocytes, is used in cancer immunotherapy to treat melanoma, non-small cell lung cancer, head and neck cancer and Hodgkin lymphoma (*Chapter 17*). Other MAbs are used as immunosuppressives following tissue/organ transplantations as outlined in *Chapter 6*.

Activation of T lymphocytes in humoral immunity

The way T lymphocytes recognise foreign epitopes is rather more complex than that used by B lymphocytes, as the receptors on T cells are unable to bind directly to epitopes on immunogenetic proteins. Instead, peptides derived from the foreign protein must be 'presented' to them by cell surface proteins encoded by genes of the **major histocompatibility complex** (MHC) found on chromosome 6 (see *Figure 6.22*). The proteins encoded by the MHC include two classes of integral membrane proteins (see *Figure 4.23*). Class I proteins are found on all nucleated cells in the body. They consist of a single polypeptide, the α chain, which is associated with a smaller protein called β_2 microglobulin (β_2M) that is not encoded by MHC genes. Class II proteins are found on only a few cell types and are made up of two polypeptides, α and β, both encoded by genes of the MHC. Both class I and II proteins have surface grooves that can bind the foreign peptide in an extended form to present it to the T lymphocyte (*Figure 4.39*). The requirement to have foreign peptides presented by different MHC molecules to T_C and T_H cells can be explained by examining their roles *in vivo*.

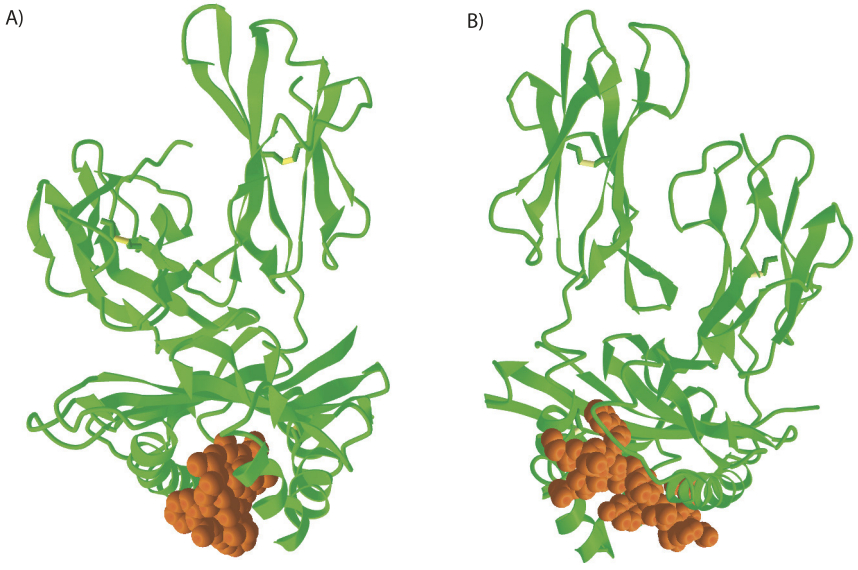


Figure 4.39 Molecular models showing (A) major histocompatibility complex (MHC) class I and (B) MHC class II molecules coloured green with the bound peptides in red. PDB files 2X4P and 1AQD, respectively.

Activation of T_C cells

The role of T_C cells is, ultimately, to destroy virus-infected cells. Since all nucleated cells are susceptible to such infections, class I MHC molecules are required by infected cells to present endogenously produced viral peptides to T_C cells. Thus, viral proteins produced within the cytoplasm of an infected cell can be hydrolysed to produce short peptides, about 8–12 amino acid residues in length. These peptides are transferred into the endoplasmic reticulum lumen where they become attached to class I proteins and the complex is then transported to the surface of the cell membrane in Golgi vesicles (*Chapter 16*).

The T_C cell is activated when the TCR recognises the complex of MHC class I protein with its bound virus-derived peptide and binds to it (*Figure 4.40*). An accessory protein, CD8, is also involved in this interaction, as it also binds to the MHC protein. Indeed, several cell adhesion molecules (see *Margin Note 5.5*) are also necessary to stabilise the binding of the two cells and to stimulate subsequent events. Following binding, the T_C cell is stimulated into a series of mitotic divisions forming a clone of cells, which then differentiate into CTLs. These cells resemble T_C cells, but the presence of vesicles containing cytotoxic proteins called granzymes and perforins (*Figure 4.41*) makes their cytoplasm appear granular. The CTL binds to a virus-infected cell using the same mechanism as the T_C cell and releases its cytotoxic proteins, which destroy the target cell. Cytotoxic T lymphocytes also secrete IL-2 and IFN-γ, which stimulate NKCs and macrophages to kill virus-infected cells.

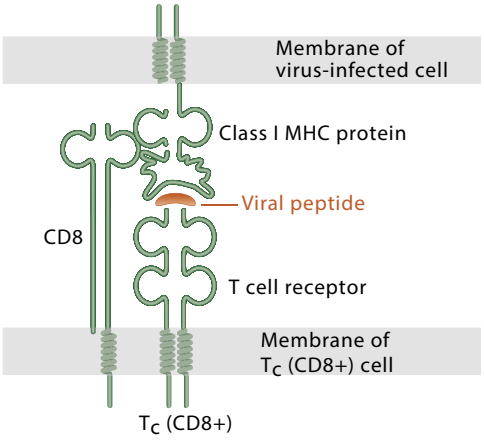


Figure 4.40 Schematic showing the interactions between a virus-infected cell and a T_C cell. MHC, major histocompatibility complex.

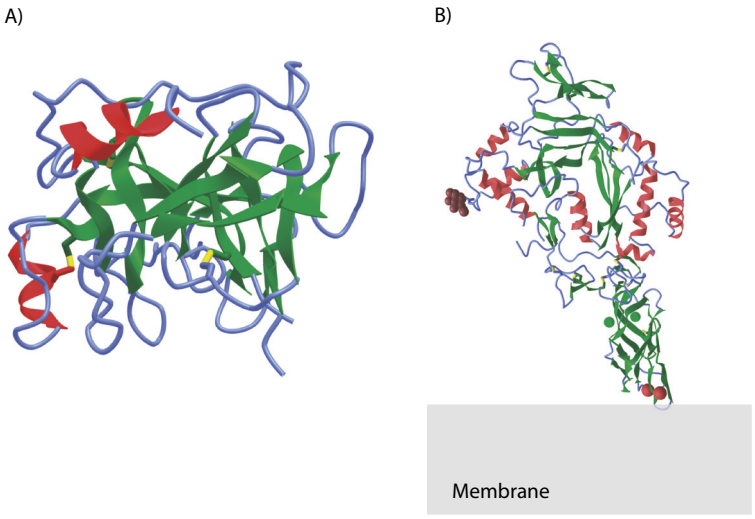


Figure 4.41 Molecular models of (A) a granzyme and (B) a perforin. PDB files 4GAW and 3NSJ, respectively.

Activation of T_H cells

When stimulated by an immunogen, TH cells secrete the cytokines required to regulate the specific and nonspecific immune responses. Thus, they are necessary for an effective immune response to the whole range of infectious agents, from viruses to multicellular parasites. To activate a T_H cell, an immunogen must be taken up by an APC. Here, its proteins are hydrolysed within endocytic vesicles and the resulting peptides of 12–19 amino acid residues are attached to MHC-encoded class II proteins and transported to the cell membrane. The MHC class II protein and foreign peptide complex is recognised and bound by a TCR (*Figure 4.42*). The coreceptor protein CD4 also binds to the MHC protein. A number of other cell membrane proteins are involved in stabilising this interaction (*Chapter 6*). Several different types of cells can act as APCs, including monocytes, macrophages, dendritic cells in the lymphoid tissues and Langerhans cells of skin. However, under certain conditions, other cells, including epithelial cells, can be induced to express MHC class II proteins, and it has been proposed that such cells may initiate autoimmune reactions (*Chapter 5*). When a T_H cell is activated by binding to the MHC class II protein–peptide complex on the surface of an APC, it is stimulated to proliferate and differentiate into cells that actively secrete cytokines. The cytokines include growth and differentiation factors that stimulate B and T cells, haematopoietic factors, and factors that stimulate mast cells, macrophages and eosinophils. There are two subsets of T_H cells: T_H1 and T_H2, which differ in the profile of cytokines they secrete, as shown in *Table 4.6*. Cytokines produced by T_H1 favour the development of cell-mediated immunity, which attacks intracellular parasites, such as viruses. Those produced by T_H2 cells stimulate humoral immunity to extracellular pathogens and parasites, including most bacteria and helminth worms and flukes. In addition, cytokines produced by T_H1 inhibit T_H2 cells and vice versa.

Margin Note 4.9 Dendritic and Langerhans cells

Dendritic cells are cells with characteristic membrane processes that visually resemble the dendrites of nerve cells. The processes give the cells a large surface area for presenting antigens to T_H cells. Dendritic cells are found in the lymphoid tissues and are highly efficient at antigen presentation. Langerhans cells in the epidermis of skin resemble dendritic cells in having multiple processes. They are involved in taking up antigens that have entered the skin before migrating to lymph nodes where they present the antigen to T lymphocytes.

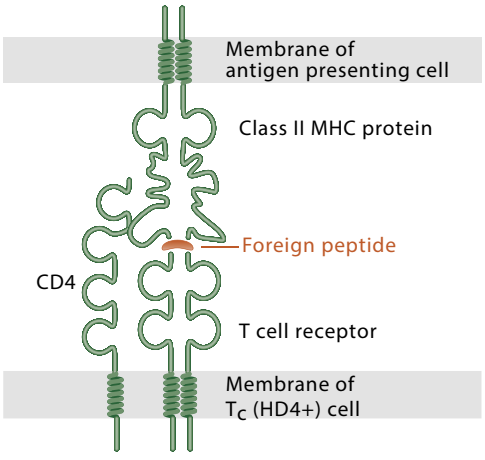


Figure 4.42 Schematic showing the activation of a T_H cell by an antigen-presenting cell. MHC, major histocompatibility complex.

T_H1 or T_H2 cytokine	Cytokine	Roles *in vivo*
T_H1	IL-2	Stimulates growth of T lymphocytes; increases activity of NKCs
T_H1	Interferon-γ	Activates macrophages and NKCs; inhibits TH2 cells
T_H1	TNF-β[a]	Promotes acute phase response; cytotoxin
T_H2	IL-4	Promotes growth and differentiation of B lymphocytes; promotes growth of TH2 cells; growth factor for mast cells
T_H2	IL-5	Promotes growth and differentiation of B lymphocytes and eosinophils
T_H2	IL-6	Promotes acute phase response; stimulates differentiation of B lymphocytes
T_H2	IL-9	Activates mast cells
T_H2	IL-10	Inhibits production of TH1 cells
T_H2	IL-13	Promotes growth and differentiation of B lymphocytes; induces IgE[b] production
T_H1 and T_H2	IL-14	B cell growth factor
T_H1 and T_H2	IL-3	Stimulates haematopoiesis
T_H1 and T_H2	GM-CSF[c]	Stimulates haematopoiesis

IL, interleukin; NKC, natural killer cell.

[a] Tumour necrosis factor. [b] Immunoglobulin E. [c] Granulocyte–macrophage colony-stimulating factor.

Table 4.6 Examples of T_H cytokines. See also *Table 4.1*. ⏎

BOX 4.5 SUPERANTIGENS AND TOXIC SHOCK SYNDROME

Superantigens, such as toxic shock syndrome toxin-1 produced by some strains of *Staphylococcus aureus* (*Chapters 2* and 3), cause massive stimulation of T_H cells. They link the T cell receptor (TCR) to the major histocompatibility complex (MHC) class II protein at sites outside the peptide-binding groove and independent of any bound peptide (*Figure 4.43*). Since this is independent of the TCR being specific to the antigen, excessive numbers of T cells are stimulated with the consequent release of high concentrations of cytokines (*Section 4.4*). This can cause a rapid onset of clinical shock with increased temperature and organ damage. Toxic shock syndrome is typically associated with localised *S. aureus* infections, such as subcutaneous abscesses or infected wounds. A number of cases have occurred with contaminated tampons, which if kept in the vagina for extended periods (more than 12 hours) provide excellent growing conditions for bacteria. Most cases are associated with teenagers but have occurred with older women. The condition requires treatment with antibiotics, particularly with the antistaphylococcals clindamycin and oxacillin, and intravenous fluids.

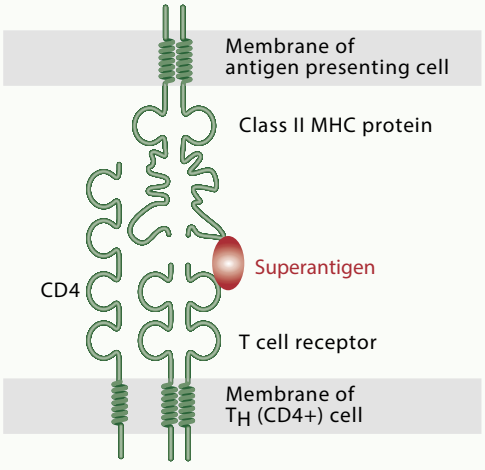


Figure 4.43 A schematic showing the linking of a major histocompatibility complex (MHC) class II molecule to a T cell receptor (TCR) by a superantigen. ⏎

Cell-mediated immunity

Cell-mediated immunity involves the direct and indirect destruction of host cells infected by viruses or other intracellular parasitic microorganisms, such as rickettsias (*Chapter 2*). Direct destruction is brought about by the production of specific cytotoxic cells capable of killing any such infected cell that induced their production. Indirect destruction is brought about by the release of cytokines that promote destruction by macrophages and LGLs. This type of immunity is the major defence against viral infections, since the destruction of virus-infected host cells prevents the replication and spread of the virus.

4.9 PARASITIC PROTOZOA, FUNGI AND HELMINTHS

Some protozoan, fungal and helminth parasites (colloquially 'worms') were described in *Chapters 2* and *3*. They all belong to the kingdom Eucarya.

Protozoan parasites often have complex life cycles, like the *Plasmodium* spp. (see *Figure 2.33*) responsible for malaria, and/or they progress through a variety of antigenic forms; for example, *Trypanosoma brucei*, the causative organism of sleeping sickness, can use 100s of variant protein coats. Such factors make mounting strong immune responses or developing viable vaccines difficult. However, protozoan parasites in the gut, blood or tissue fluid can promote humoral responses, although the presence of the parasites at these sites is likely to be transient. Protozoa with an intracellular life cycle stage may be attacked by a cell-mediated response.

Fungal infections or mycoses are caused by several different groups of fungi. They often occur in immunocompromised individuals, for example oral and vaginal *Candida albic*ans infections in primary and secondary immunodeficiency diseases (*Chapter 3*). Cutaneous infections, such as ringworm and athlete's foot, are relatively common. Most are cleared quickly in healthy individuals. Subcutaneous infection may occur following trauma and can be more serious, resulting in inflammation that is sufficiently severe to cause tissue damage. Abdominal organs, bones, the central nervous system and the lungs can all be infected with so-called deep mycoses.

The major defence against mycoses is innate response. Fungal cell walls contain chitin, a polymer of *N*-acetylglucosamine residues, and β glucans and mannans. The residue components of the latter two are, as may be expected from their names, β glucose and mannose. These can constitute effective PAMPs (*Section 4.5*) on the fungal pathogens. Their recognition by PRRs allows the mobilisation of, for example, complement.

Unfortunately, helminth worms generally promote fairly weak immune responses. Most are gut parasites entering the patient in contaminated water or food. Schistosomes are an exception to this and penetrate the skin. These are exclusively extracellular parasites and can stimulate an IgE-mediated response, as mentioned earlier. Tuft cells in the gut epithelium sense the presence of worms and secrete **alarmins**, for example IL-25 (also called IL-17E; *Figure 4.44*). Alarmins induce immune cells to release type 2 cytokines: IL-4, IL-5 and IL-13. The cytokines IL-4 and IL-5 promote the secretion of worm-specific IgE from B cells, which binds to the surface of the worm and recruits eosinophils to the site. Here, their release of toxic mediators on to the worm may lead to its expulsion from the gut or can even be fatal to it. The release of IL-13 stimulates gut goblet cells to increase mucus production (see *Margin Note 4.1*), which also helps in the physical clearance of the worm. The release of leukotriene fatty acids (*Chapters 5* and 10) and histamine from mast cells stimulates strong sporadic muscular contractions that can cause projectile vomiting and explosive diarrhoea, which may promote the expulsion of the parasites from the body.

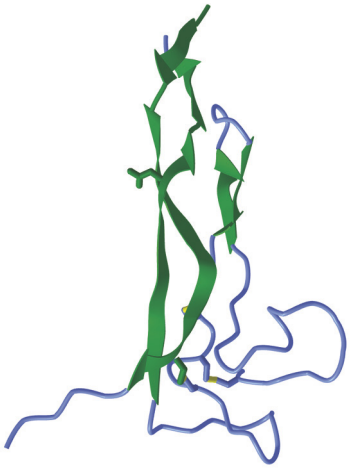


Figure 4.44 Molecular model of the alarmin interleukin-25 (IL-17E). PDB file 7UWL. ⏎

4.10 CASE STUDIES

CASE STUDY 4.1

Maria, a 28-year-old biology teacher, suspected she might be pregnant, as it was eight weeks since the start of her last period. Her pregnancy was confirmed when she carried out a home pregnancy test. Maria and her partner were delighted and began to make plans for the future. However, Maria became concerned when, during the eighth week of her pregnancy, several children in one of her classes developed German measles (rubella). Maria knew that, if she developed rubella, there was a high risk that the baby would be harmed. Unfortunately, she could not remember if or when she had ever been immunised against rubella. She consulted her family doctor who took a blood sample and had it tested for antibodies. The test revealed that Maria's blood was positive for IgM antibodies specific for rubella but rubella-specific IgG antibodies were not detected.

Questions

(a) What are the consequences of this result for the unborn baby?

(b) Why would the consequences be different if IgG antibodies specific for rubella had been detected?

(c) What counselling and/or advice would you recommend to Maria?

CASE STUDY 4.2

Alfred is a 70-year-old man who is suspected of having a myeloma, that is, a plasma cell tumour.

Question

Suggest tests that could be carried out to confirm the diagnosis. Assuming the plasma cell was producing IgG, suggest an assay that could be used to measure the concentration of IgG in his blood.

CASE STUDY 4.3

Thea, a healthy 17-year-old sixth form college student, suddenly became ill with vomiting, dizziness, increased temperature and generalised body aches. She was quickly taken to the accident and emergency department where her condition immediately raised concern. She developed a bright red rash that quickly enveloped most of her body. Although listless, she did not lose consciousness and was able to answer questions. She had not abused alcohol or drugs or been in contact with any ill person. Her recent history revealed she was menstruating and had inserted a tampon the previous day. She had worked late on a college set essay, slept late and began to feel ill almost immediately after waking. A clinical examination showed (reference values in parentheses):

Temperature	39.6 (normal 37°C)
Pulse	148 (normal 72/min)
Blood pressure	100/66 (normal 120/70 mmHg)
Respiratory rate	28 (normal 18/min)

Laboratory tests showed a leukocyte count of 21.5×10^9/L (normal $4–11\times10^9$/L), with neutrophils predominating. Blood, urine, and cerebrospinal and vaginal fluids were taken and subjected to microbiological examination.

Thea was eventually diagnosed with toxic shock syndrome and appropriately treated. Fortunately, she made a complete recovery.

Questions

(a) What is the most likely source of Thea's infection?
(b) What is the significance of the high leukocyte count and the preponderance of neutrophils?
(c) What would have been the most likely results of the microbiological examinations of the body fluids?

4.11 SUMMARY

The basis of the actions of the immune system is its ability to distinguish self from nonself. It defends the body in a variety of nonspecific and specific ways. Nonspecific defences include structural barriers and complement. Nonspecific responses include inflammation and the acute phase response, while specific responses include the production of antibodies, known as humoral immunity, and the production of cytotoxic cells in cell-mediated immunity. Specific defences involve the development of immune responses against infectious agents. An effective immune response results from the complex interaction between nonspecific and specific cells, which are coordinated, controlled and regulated by signalling molecules called cytokines. The nonspecific cells include the monocytes, LGLs and PMNs found in the blood. The specific cells are the small lymphocytes found in the blood and lymphoid tissues. Humoral immunity is effective at dealing with extracellular bacteria and multicellular parasites, while cell-mediated immunity is effective at killing virus-infected cells. Small lymphocytes are highly specific, since they bear cell surface receptors for epitopes found on foreign proteins. Small lymphocytes belong to one of two major subsets: the B lymphocytes responsible for humoral immunity and the T lymphocytes, some of which, the T_C cells, can develop into cytotoxic cells, while others, the T_H cells, regulate immune responses through the secretion of cytokines.

Immunoglobulins are glycoproteins that are highly specific for epitopes. This specificity has enabled them to be used in the detection and quantitation of antigens in highly sensitive techniques such as ELISA.

QUESTIONS

1. Which of the following is **NOT** part of the nonspecific defence against invading microorganisms?
(a) Skin.
(b) Mucus.
(c) Antibodies.
(d) Lysozyme.
(e) Lactic acid.

2. Which of the following is **NOT** found in the blood?
(a) Macrophage.
(b) Monocyte.
(c) Neutrophil.
(d) Small lymphocyte.
(e) LGL.

3. Which of the following statements best describes complement?
(a) A defence protein found in sweat.
(b) A group of proteins that results in the lysis of bacteria.
(c) A cytokine produced by macrophages.
(d) An antibody that lyses bacteria.
(e) A protein found in the cell wall of bacteria.

4. Pair up the cells in column 1 with the appropriate cell surface molecule in column 2.

Column 1	**Column 2**
All mature T lymphocytes	CD4+
All B lymphocytes	CD8+
Macrophages	CD3
Helper T lymphocytes	Surface Ig
Cytotoxic T lymphocytes	Receptors for complement

5. Which Ig class in each case best fits the following description?
(a) The only antibody that crosses the placenta.
(b) The most abundant antibody in the blood.
(c) Also known as the secretory antibody.
(d) Produced in response to parasitic worms.
(e) The antibody with the largest M_r.

6. Why is it unnecessary for erythrocytes to have MHC-encoded class I proteins?

7. Use *Figure 4.11(B)* to determine the approximate diameter of the pore of a MAC.

8. Arrange the two following lists into their most appropriate pairings.

Ibuprofen	Class 1 cytokine
Lysozyme	Alternative complement pathway
GM-CSF	TNF cytokine
IFNs	Chemokine
BAFF	Type II IFN
IL-8	C5 convertase
C3bBb	Mild anti-inflammatory treatment
C3(H_2O)	Antimicrobial enzyme
IFN-γ	Class 2 cytokines
C4b2a3b	C3 convertase

9. State whether the following statements are TRUE or FALSE.
(a) Cytokines bind to intracellular receptors.
(b) The IL-1 family are chemokines secreted in response to viral infections.
(c) IL-8 is a chemokine.
(d) Class 1 cytokines are the largest family of chemokines.
(e) Class 1 cytokine molecules have a common structure consisting of a four-α-helical motif.
(f) Complement functions in the clearance of immune complexes.
(g) Properdin is a complement protein.
(h) The classical complement pathway can be activated by IgG and IgE.
(i) The hydrolysis of an internal ester bond converts C3 to C3(H_2O).
(j) Opsonisation can be effected by some types of neutrophils.

10. Which of the following is the odd ONE out?
(a) C4a.
(b) C3bBb.
(c) MASP-3.
(d) C4b2a3b.
(e) C5bC6C7C8$(C9)_n$.

FURTHER READING

Bajic G *et al.* (2015) Complement activation, regulation, and molecular basis for complement-related diseases. *EMBO J* 34:2735–2757.

Bournazos S *et al.* (2017) Signaling by antibodies: recent progress. *Annu Rev Immunol* **35**:285–311.

Chiu ML *et al.* (2019) Antibody structure and function: the basis for engineering therapeutics. *Antibodies (Basel)* **8**:55.

Clark IA and Vissel B (2017) The meteorology of cytokine storms, and the clinical usefulness of this knowledge. *Semin Immunopathol* **39**:505–516.

Custodio H (2016) Protozoan parasites. *Pediatr Rev* **37**:59–69.

Fajenbaum DC and June CH (2020) Cytokine storm. *N Engl J Med* **383**:2255–2273.

Filardy AA *et al.* (2018) Human kinetoplastid protozoan infections: where are we going next? *Front Immunol* **9**:Article 1493.

Filbey KJ *et al.* (2021) Editorial: Recent advances in the immunology of helminth infection – protection, pathogenesis and panaceas. *Front Immunol* **12**:Article 663753.

France MM and Turner JR (2017) The mucosal barrier at a glance. *J Cell Sci* **130**:307–314.

Grumach AS and Kirschfink M (2014) Are complement deficiencies really rare? Overview on prevalence, clinical importance and modern diagnostic approach. *Mol Immunol* **61**:110–117.

Heesters BA *et al.* (2016) Antigen presentation to B cells. *Trends Immunol* **37**:844–854.

Hidalgo A *et al.* (2019) The neutrophil life cycle. *Trends Immunol* **40**:584–597.

Hunter CA and Jones SA (2015) IL-6 as a keystone cytokine in health and disease. *Nat Immunol* **16**:448–457.

Kalliolias GD and Ivashkiv LB (2016) TNF biology, pathogenic mechanisms and emerging therapeutic strategies. *Nat Rev Rheumatol* **12**:49–62.

Klenerman P (2017) *The Immune System: A Very Short Introduction*. Oxford University Press.

Kobayashi T and Mattarollo SR (2019) Natural killer cell metabolism. *Mol Immunol* **115**:3–11.

Lang R and Schick J (2017) Review: Impact of helminth infection on antimycobacterial immunity – a focus on the macrophage. *Front Immunol* **8**:Article 1864.

Lu F and Huang S (2017) The roles of mast cells in parasitic protozoan infections. *Front Immunol* **8**:Article 363.

Lu R-M *et al.* (2020) Development of therapeutic antibodies for the treatment of diseases. *J Biomed Sci* **27**:1.

Marshall JS *et al.* (2018) An introduction to immunology and immunopathology. *Allergy Asthma Clin Immunol* **14**:49.

Morgan BP *et al.* (2017) Molecular cell biology of complement membrane attack. *Semin Cell Dev Biol* **72**:124–132.

Pathakumari B *et al.* (2020) Immune defence to invasive fungal infections: a comprehensive review. *Biomed Pharmacother* **130**:Article 110550.

Qidwai T (2016) Chemokine genetic polymorphism in human health and disease. *Immunol Lett* **176**:128–138.

Ragland SA and Criss AK (2017) From bacterial killing to immune modulation: recent insights into the functions of lysozyme. *PLoS Pathog* **13**(9):e1006512.

Rajasekaran S *et al.* (2017) TLR specific immune responses against helminth infections. *J Parasitol Res* **2017**:6865789. doi: 10.1155/2017/6865789.

Schroder K and Tschopp J (2010) The inflammasomes. *Cell* **140**:821–832.

Souza-Fonseca-Guimaraes F *et al.* (2019) The emergence of natural killer cells as a major target in cancer immunology. *Trends Immunol* **40**:142–158.

Veliz K and Beier UH (2021) Transcriptional regulation of T cell metabolism and metabolic control of T cell gene expression. *Curr Opin Genet Dev* **70**:83–88.

Wang Q *et al.* (2022) Molecular biology techniques. In: Ahmed N, Hedley G and Wang Q (eds), *Biomedical Science Practice*, 3rd edn. Oxford University Press.

Yamanaka YJ *et al.* (2913) The dynamic lives of T cells: new approaches and themes. *Trends Immunol*, **34**:59–66.

Yates C and Wang Q (2022) Immunological techniques. In: Ahmed N, Hedley G and Wang Q (eds), *Biomedical Science Practice*, 3rd edn. Oxford University Press.

Useful website

www.iuphar.org

Disorders of the immune system

5

OBJECTIVES

After studying this chapter you should be able to:

- describe some features of selected
  - immunodeficiency diseases
  - autoimmune diseases
  - immunological hypersensitivities
- outline ways in which immunological disorders are investigated and treated

5.1 INTRODUCTION

A fully functioning immune system is essential for maintaining health but, like all body systems, is itself subject to many clinical disorders. The first are **immunodeficiency diseases** that result from the absence or malfunction of one or more immune system components, making patients more susceptible to infections and even to certain types of cancers (*Chapter 17*). Second are **autoimmune diseases**, which occur when the immune system fails to recognise 'self' components and mounts an attack against the patient. Third, **immunological hypersensitivities** occur when immune responses, often mounted against seemingly innocuous immunogens, cause tissue damage.

5.2 IMMUNODEFICIENCY DISEASES

Inborn errors of immunity (IEIs), previously called primary immunodeficiency diseases (PIDs), are rare disorders resulting from direct defects in proteins, cells or organs of the immune system. All IEIs are inherited or have a genetic component. Most, but not all, are congenital, although some do not manifest themselves until later in life.

Patients with immunodeficiency diseases invariably suffer infections more frequently and with increased severity than otherwise healthy individuals. The types of infections depend largely upon the nature of the immune deficiency but some general rules to guide diagnoses are possible. For example, a deficiency in antibodies generally causes recurrent bacterial infections of the chest and boils in the skin. Deficiencies of T lymphocytes can result in viral and intracellular bacterial infections, with, for instance, herpes simplex viruses (*Table 3.1*) and mycobacteria (*Chapter 3*), respectively. Fungal pathogens like *Pneumocystis carinii* and opportunistic agents such as *Candida albicans* also cause clinical problems. Deficiencies in complement proteins may result in recurrent meningococcal meningitis, bacterial pneumonia and septicaemia, while defective phagocytic cells can result in boils and abscesses on the skin.

Disorders may also arise because of an IEI. Thus, absent or defective complement proteins may be associated with systemic lupus erythematosus (SLE), an autoimmune disorder (*Section 5.4*). Certain types of cancers, such as non-Hodgkin lymphoma and Kaposi sarcoma, are also more frequent in immunodeficient patients because they are linked to infections with certain viruses (*Chapter 17*). The former is a tumour of the lymph glands, while the latter is derived from blood vessels but frequently shows in the skin (*Figure 5.1*).

Secondary immunodeficiency diseases (SIDs) arise because of other conditions. For example, some viruses are associated with an immunosuppression, which may be transient, such as in measles, or permanent, as in human immunodeficiency virus (HIV) infection. Some conditions that can give rise to SIDs are listed in *Table 5.1*. This chapter will concentrate on IEIs.

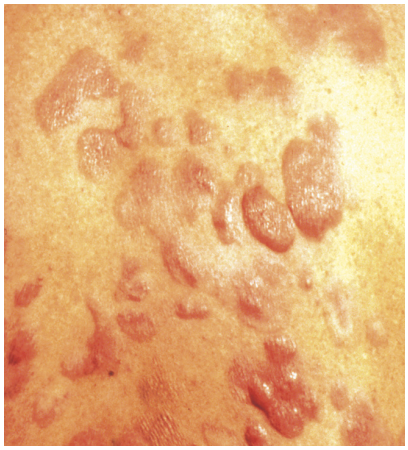


Figure 5.1 Skin of the back of a patient suffering from Kaposi sarcoma. Courtesy of National Cancer Institute, USA. ⏎

DOI: 10.1201/9781003296836-5

Contributing factor	Comments
Malnutrition	A correctly functioning immune system requires essential nutrients (*Chapter 10*)
Physical trauma	For example, extensive burns and surgery are associated with immunosuppression
Immunosuppressive drugs	Drugs given to prevent transplant rejection (*Chapter 6*), chemotherapy for cancer (*Chapter 17*)
Infectious diseases	For example, HIV/AIDS (*Chapter 2*), measles, cytomegalovirus infection
Haematological disease	For example, myeloma (*Chapter 4*), leukaemia (*Chapter 17*)
Immature immune system	For example, in premature infants
Ageing	All body systems malfunction with age; the immune system is no different (*Chapter 18*)
Metabolic and hereditary disorders	For example, diabetes mellitus (*Chapter 7*)
Stress	Stress hormones, such as cortisol (*Chapter 7*), can be immunosuppressive

HIV/AIDS, human immunodeficiency virus/acquired immunodeficiency syndrome.

Table 5.1 Some factors that contribute to secondary immunodeficiency diseases (SIDS). ⏎

5.3 PRIMARY IMMUNODEFICIENCY DISEASES

The IEI Committee (which, prior to 2017, was known as the Primary Immunodeficiency Diseases Expert Committee) of the International Union of Immunological Societies (IUIS) lists and classifies IEIs (*Table 5.2*) on a regular basis. In 2024, it recorded 555 inherited IEIs.

IUIS group	IEI category
I	Immunodeficiencies affecting cellular and humoral immunity
II	Combined immunodeficiencies (CIDs) with associated syndromic features
III	Predominantly antibody deficiencies
IV	Diseases of immune dysregulation
V	Congenital defects of phagocytes
VI	Defects in intrinsic and innate immunity
VII	Autoinflammatory disorders
VIII	Complement immunodeficiencies
IX	Bone marrow failure
X	Phenocopies of IEIs

Table 5.2 The International Union of Immunological Societies (IUIS) classification of inborn errors of immunity (IEIs). ⏎

Obviously, this chapter cannot deal with all of them! Rather, the following IEIs will be highlighted to illustrate their range:

- combined T and B lymphocyte deficiency (severe combined immunodeficiency, SCID)
- selected antibody deficiencies
- DiGeorge anomaly

- Disorders related to the Wiskott–Aldrich gene (*WAS*)
  - Wiskott–Aldrich syndrome (WAS)
  - X-linked thrombocytopenia
  - X-linked congenital neutropenia
- Chédiak–Higashi syndrome
- deficiencies of phagocyte functioning
  - chronic granulomatous disease
  - leukocyte adhesion deficiencies
- antibody deficiencies
  - hypogammaglobulinaemias
  - agammaglobulinaemias
- complement deficiencies

Combined T and B lymphocyte deficiencies

Deficiencies affecting both B and T lymphocyte (*Chapter 4*) numbers and/or their functions, termed **severe combined immunodeficiency** (SCID), are life threatening. They occur with a frequency of approximately 1 in 40 000–100 000 births. This group of disorders are associated with more than 20 different mutations. Some are inherited in an X-linked fashion (*Chapter 15*), and hence approximately three times more boys than girls are affected, but others are autosomal recessive traits. Both types result in the incomplete development of T cells with or without B cells.

Depending on the mutation(s) involved, T and B cells may both be decreased (T^-B^- SCID) or B cells numbers may be normal or increased (T^-B^+ SCID). In the latter, the absence of T cells renders B cells functionally inactive because they require cytokines produced by T_H cells to stimulate the production of antibodies (*Chapter 4*). These phenotypes may be subdivided depending on the presence/absence of natural killer cells (NKCs).

The disorder becomes apparent in the first three to six months of life. It is characterised by chronic viral, bacterial, protozoan and fungal infections, constant diarrhoea, and a general failure by the baby to thrive. If given live viral vaccines, such as the measles, mumps and rubella (MMR) and polio vaccines (*Chapter 3*), affected infants may suffer generalised viral infections. Laboratory tests show leukocyte counts less than 3000 per μL (reference range for an infant of six months is 4000 to 13 500 per μL). The lymphocytes present are functionally inactive and do not respond *in vitro* to known **mitogens** (*Box 5.1*). Chest X-rays show an abnormally small or absent thymus.

BOX 5.1 LABORATORY TESTS FOR LYMPHOCYTE FUNCTION ⏎

Tests are available to assess lymphocyte functions. For example, small lymphocytes can be incubated with antigens to which a child has been exposed. If the child is immune, the cells will respond by proliferating. The extent of cell growth can be determined by the amount of radiolabelled thymidine incorporated into the DNA of the dividing cells. The antigen will stimulate only those cells that bear specific receptors, perhaps only 1 in 100 000, and so the test may be relatively insensitive. However, some proteins and other types of molecules, such as lipopolysaccharides, are mitogens and act as **polyclonal** activators because they stimulate many different clones of lymphocytes to divide. Lectins are mitogens that bind to carbohydrate residues on the lymphocyte plasma membrane and stimulate lymphocyte proliferation. Lectins are often isolated from plants (*Table 5.3*) although some are obtained from other organisms. It is also possible to stimulate lymphocytes to divide by incubating them with antibodies directed at cell surface receptors. Thus, an antibody to CD3 will stimulate all mature T lymphocytes to divide. A typical response of T lymphocytes from a healthy individual to phytohaemagglutinin (PHA) is shown in *Figure 5.2*. A failure to respond in this way is indicative of a lymphocyte defect. In patients with severe combined immunodeficiency (SCID), the proliferative response is less than 10% of control values.

Mitogen	Derived from	Small lymphocytes stimulated
Phytohaemagglutinin (PHA)	*Phaseolus vulgaris*	T lymphocytes
Concanavalin A (ConA)	*Canavalia ensiformis*	T lymphocytes
Poke weed mitogen (PWM)	*Phytolacca americana*	B lymphocytes
Lipopolysaccharide (LPS)	*Escherichia coli*	B lymphocytes
Anti-CD3	Monoclonal antibody originating from mice	T lymphocytes
Anti-immunoglobulin	Monoclonal antibody originating from mice	B lymphocytes

Table 5.3 Polyclonal activators of small lymphocytes. ⏎

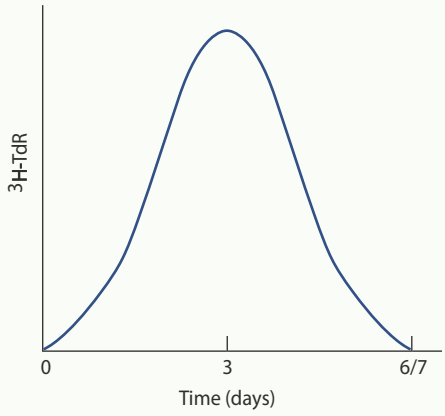


Figure 5.2 The stimulation by phytohemagglutinin (PHA) on all T cells as measured by the uptake of tritiated thymidine (^{3}H-TdR) by the dividing cells.

Severe combined immunodeficiencies are caused by a whole raft of mutations. Approximately 20% of $T^-B^-NK^-$ SCID arises from mutations in *ADA* located at chromosome 20q13.12 (*Chapter 15*), which encodes adenosine deaminase (ADA; *Figure 5.3*). Its absence results in the accumulation of adenosine and related metabolites that are toxic to small lymphocytes and result in their death by apoptosis (*Chapters 15* and *17*). A similar syndrome arises from mutations in *PNP* leading to a deficiency of purine nucleoside phosphorylase (PNP; *Figure 5.4*).

The Artemis gene, *DCLRE1C*, encodes a nuclease essential for antibody production. Like *ADA*, it results in $T^-B^-NK^-$ SCID. The *RAG1* and *RAG2* genes encode enzymes that are also essential for recombination and antibody production. Mutations result in $T^-B^-NK^+$ SCID. Mutations in *ADA*, *PNP*, *DCLREIC* and *RAG1* and *2* resulting in SCID are all inherited in an autosomal recessive fashion. Autosomal recessive mutations in genes encoding a variety of CD proteins also lead to molecular lesions and result in SCID phenotypes. However, approximately 40% of T^-B^+ SCID cases are X-linked and arise from mutations in genes for interleukin (IL) receptor proteins, so that cells are unable to respond to, for example, IL-2, -4, -7, -9 and -15.

If not diagnosed early and treated appropriately, children with SCID usually die from infections within their first year. Management of this condition includes taking measures to avoid infection and administering antiviral, antibacterial and antifungal drugs. Keeping such infants in totally aseptic conditions for extended periods is neither feasible nor ethical, since this would preclude direct contact with other humans. Their immune system may be restored with a bone marrow transplant or haematopoietic stem cell transplantation (BMT and HSCT, respectively; see *Box 5.2*). This can lead to long-term survival but is not without danger, chiefly graft versus host disease (GVHD; *Chapter 6*). Gene therapy is being explored with some success.

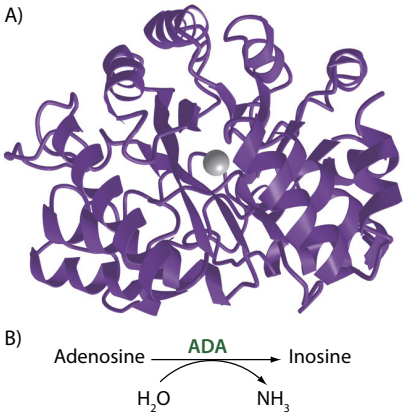


Figure 5.3 (A) Molecular model of adenosine deaminase (ADA); its Zn^{2+} is depicted as a grey sphere. PDB file 3IAR. (B) The reaction catalysed by ADA.

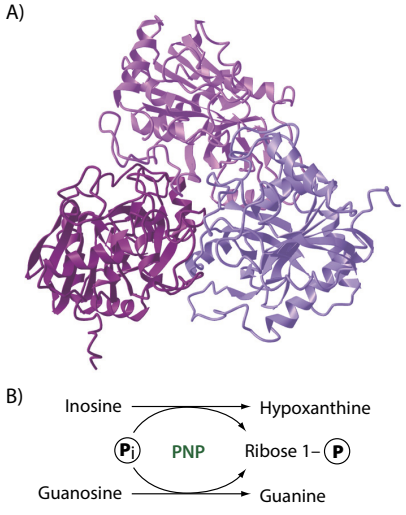


Figure 5.4 (A) Molecular model of purine nucleoside phosphorylase (PNP). PDB file 1M73. (B) Examples of reactions catalysed by PNP.

BOX 5.2 BONE MARROW TRANSPLANT/HAEMATOPOIETIC STEM CELL TRANSPLANTATION: WHAT'S THE DIFFERENCE?

Haematopoietic stem cells (HSCs, *Chapter 13*) are found in bone marrow, vascular blood and umbilical cords. A **stem cell transplant** (see *Chapter 6*) can refer to a bone marrow transplant (BMT), a peripheral blood stem cell transplant (PBSCT) or an umbilical cord blood transplantation (UCBT). The only difference between them is the *source* of the stem cells to be transplanted.

Collecting bone marrow requires the donor to be under general anaesthesia. A syringe is then used to draw bone marrow from the interior of the iliac crest (hip bone). The process takes up to two hours. Donors generally experience back ache for one to three days after the procedure, which usually responds to proprietary painkillers. Full recovery takes two to five days.

Circulating blood contains fewer HSCs than bone marrow but its numbers can be increased prior to collection by stimulating the marrow with recombinant granulocyte colony-stimulating factor (GCSF, *Figure 5.5*), which leads to increased numbers of stem cells in the blood. The cells are harvested by **apheresis**, a term derived from the Greek meaning 'to carry away' (see

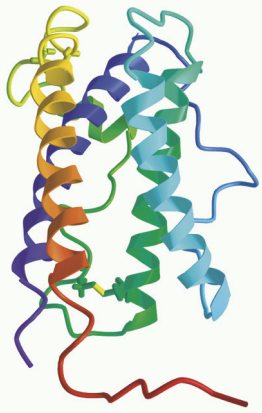


Figure 5.5 Molecular model of recombinant granulocyte colony-stimulating factor (GCSF). PDB file 1GNC.

Chapter 6). A needle inserted in an arm of the awake donor allows his/her blood to flow through a tube to an 'apheresis machine'. Here the donor blood is separated into its component parts by either centrifugation or membrane filtration. The former method is preferred for harvesting specific types of blood cells. The stem cells are collected into a small bag and the remaining blood returned to the donor via a tube to the other arm. The process takes four to six hours, depending on the number of cells required. Following the procedure, most donors feel tired but fully recover within 24–48 hours.

Whatever the method of collection, the transplanted stem cells are infused into the patient's bloodstream. From there, the stem cells migrate into the largest bones and, if the transplant is successful, become established and restore immune system functions.

Approximately 80% of stem cell transplants are PBSCTs, having largely replaced BMTs because:

- the method of collecting stem cells is easier
- normally more cells can be obtained
- donors suffer less pain and inconvenience
- neutrophil and platelet functions recover more rapidly with a faster immune reconstitution in the patient

Blood from umbilical cords contains insufficient cells for use with adult patients and therefore its use is restricted to transplants involving children.

Antibody deficiencies

Deficiencies involving immunoglobulins (Igs) of all classes are commonly referred to as **hypogammaglobulinaemias** or **agammaglobulinaemias**, which depends on the level of deficiency (*Table 5.4*). However, with some disorders there may be a selective deficiency of a single immunoglobulin class, as in selective IgA deficiency, or a dysregulation where some antibody classes are reduced while others are increased.

Transient hypogammaglobulinaemia (TH) occurs when the start of production of IgG in babies is delayed. The incidence of TH has been estimated as 23–61 per million births. It possibly arises when the maturation of helper T lymphocytes is itself delayed. During pregnancy, IgG is the only immunoglobulin to cross the placenta from mother to fetus. Thus, a newborn baby has adult concentrations of IgG, most of which is maternally derived (*Figure 5.6*). During the first three months postpartum, the maternally derived IgG is catabolised and steadily disappears from the baby's circulation. Between three and six months, serum IgG concentrations may be quite low, after which their biosynthesis begins to increase, and should normally attain 'adult' values at 12–18 months of

Immunoglobulin (Ig)	Serum concentration in adults (g/L)
IgM	0.5–2.0
IgG	7.2–19.0
IgA	0.8–5.0
IgD	Trace
IgE	Trace

Table 5.4 The normal (physiological) adult concentration ranges of the five classes of immunoglobulins.

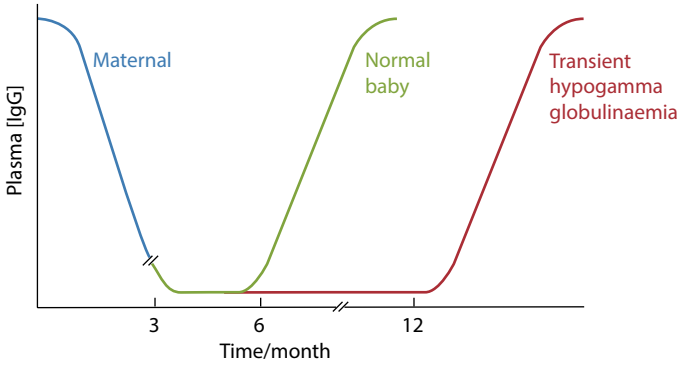


Figure 5.6 Transient hypogammaglobulinaemia (TH). The production of IgG in the baby normally starts at three to six months of age, by which time the maternal IgG in the baby has almost disappeared. In TH, the production of IgG by the baby is delayed.

age. In infants with TH, production of IgG is delayed considerably, sometimes for as long as two years. During this time, the infant is susceptible to recurrent infections with **pyogenic** (pus-producing) bacteria and antibiotics must be administered.

Selective IgA deficiency, as its name implies, affects only this single class of immunoglobulin. This is a relatively common disorder with an incidence of 1 in 500–700 of the White population, although the frequency is much lower in other ethnic groups. Deficiencies in IgM or IgG are also known but are much rarer. Many IgA-deficient individuals are asymptomatic, with the condition only being detected during investigations of other disorders. In contrast, other patients with selective IgA deficiency suffer recurrent infections, typically ear infections, sinusitis and pneumonia. A high proportion of sufferers also develop autoantibodies that are directed against a variety of self antigens. Approximately a third present with autoimmune diseases, such as SLE (*Section 5.4*). It is not known which features determine the severity of the disease.

A baby or infant who presents with a history of recurrent infection, chronic diarrhoea and autoimmune disease should be investigated for selective IgA deficiency. This can be confirmed by measuring serum immunoglobulin concentrations. Values of IgA less than 0.07 mg/L would confirm the deficiency. Other immunoglobulins could be within their usual ranges, but values of IgM may even be above normal because some patients show compensatory increases. Treatment of selective IgA deficiency normally involves using antibiotics to treat bacterial infections. Immunoglobulin replacement therapy (IRT; see below) is not usually necessary. If the disease presents with autoimmunity then anti-inflammatory drugs, such as corticosteroids, may be given. The prognosis is good, with patients living normal lifespans. However, approximately 10% of patients with a selective IgA deficiency also have a deficiency of the IgG_2 subclass, which is usually produced in response to polysaccharide antigens. Patients with both defects suffer more severe bacterial infections, especially by encapsulated bacteria, and IRT may be appropriate in these cases. Patients should also be monitored for progression to common variable immunodeficiency (CVID).

Common variable immunodeficiency is the most common IEI with an incidence of 1 in 10 000–50 000. Some cases are diagnosed at one to five years, but many not until early adulthood. The cause(s) of CVID is (are) unknown; most cases are sporadic. However, this heterogeneous group of disorders appears to have a genetic basis. The majority of identified mutations encode proteins of immune cells, which is, perhaps, not unexpected.

Most CVID patients show low blood concentration of IgG and IgA, with near normal or 50% of normal values of IgM, and normal lymphocyte counts. The latter allows CVID to be distinguished from other antibody deficiencies such as X-linked agammaglobulinaemia (see below). Some patients also have impaired cell-mediated immunity. Unfortunately, B lymphocytes are immature and, when stimulated, do not differentiate into antibody-secreting plasma cells following the binding of an antigen (*Chapter 4*) owing to defects in their cell surface receptors or signal transduction mechanisms (*Chapter 7*). However, it is possible that in some patients there may be other defects, such as mutations in immunoglobulin regulatory genes. In addition, many CVID patients have defective CD4+ T lymphocytes so that the T and B cell interactions required for antibody production are impaired. Indeed, T cell defects have been reported in approximately 30% of CVID patients.

Patients with CVID present with recurrent pyogenic infections, especially of the respiratory tract, commonly involving *Streptococcus pneumoniae*, *Staphylococcus aureus*, *Haemophilus influenzae* and *Moraxella catarrhalis*. This can lead to **bronchiectasis** (abnormally dilated bronchi and bronchioles). Those patients who also have some impairment in cell-mediated immunity also suffer infections with mycobacteria and the fungus *P. carinii*. Such patients also suffer recurrent severe infections with some herpes viruses (*Chapter 3*) and may develop viral illness when immunised with live viral vaccines.

Diagnosis of CVID relies on measurements of immunoglobulins, including specific antibodies to common vaccines, and on ruling out other immunodeficiencies, such as X-linked agammaglobulinaemia. Treatment involves IRT using pooled immunoglobulins obtained from healthy donors. All pooled immunoglobulin preparations are treated to inactivate viruses, such as hepatitis virus or HIV, that may be present. These preparations are available commercially and need to be given intravenously every three to four weeks to maintain plasma concentrations and protect against infections. The dose depends on the weight of the patient and is usually 400–600 mg kg^{-1}. Weekly subcutaneous lower doses can also be administered. These can be given at home, which may be more convenient for the patient. Intramuscular injections that administer greater volumes than subcutaneous doses can also be given.

X-linked agammaglobulinaemia (XLA) or Bruton disease is caused by a deficiency in Bruton tyrosine kinase (BTK; *Figure 5.7*). This is a rare inherited disorder with an incidence of about 1 in 250 000 males. The deficiency of BTK is due to one or more of 500

Margin Note 5.1
Bruton disease/X-linked agammaglobulinaemia

In 1952, the physician Ogden Bruton (1908–2003) reported the case of an eight-year-old boy who presented with multiple pyogenic infections, repeated bouts of pneumonia with the same serotype of *Pneumococcus* and repeated attacks of mumps. Electrophoresis of his serum showed it to lack globulins, hence the term **agammaglobulinaemia**. Subsequent gene mapping over the years localised the genetic lesion to the X chromosome, giving the name *X-linked agammaglobulinaemia* (XLA). In 1993, the disease was shown to arise from mutations in the Bruton tyrosine kinase gene (*BTK*).

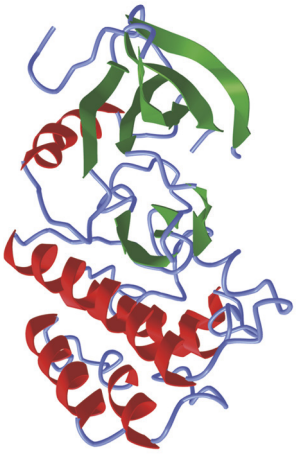


Figure 5.7 Molecular model of Bruton tyrosine kinase (BTK). PDB file 3GEN.

different mutations in *BTK* located on the long arm of the X chromosome. Bruton tyrosine kinase activity is involved in several intracellular signalling pathways (*Chapter 7*) and is required for the maturation of pre-B cells to B lymphocytes in the bone marrow. In XLA, B lymphocytes therefore account for less than 2% of circulating lymphocytes, and plasma cells are not present in the spleen and lymph nodes. The concentrations of all classes of immunoglobulin in the blood are extremely low. However, circulating T lymphocytes are normal. Tonsils and adenoids may be absent, as demonstrated by radiography.

BOX 5.3 HYPERIMMUNOGLOBULIN SYNDROMES

Hyper-IgE, or Buckley–Job, syndrome is caused by autosomal dominant mutations in *STAT3*, located within chromosome region 17q21.2. This gene encodes the signal transducer and activator of transcription 3 (STAT3) protein, which is necessary for the intracellular signalling pathways associated with several cytokines. Mutations in *STAT3* therefore limit responses to a variety of antigenic stimuli. This, in turn, reduces the numbers of some types of T_H cells that are particularly necessary to combat bacterial and fungal infections, for example *Staphylococcus aureus* and *Candida albicans*. However, the reasons why this increases the concentration of IgE are unclear. Symptoms of hyper-IgE syndrome include recurrent pneumonia, skin and lung abscesses, facial abnormalities, and fragile bones and joint lesions.

Hyper-IgD (*Figure 5.8*) or periodic fever syndrome (HIDS) is a rare autoinflammatory disease first reported in 1984. The disease is inherited as an autosomal recessive trait following mutations in *MKD* sited at 12q24.11. This gene encodes mevalonate kinase (MK); hence, a more accurate term for the syndrome may be *mevalonate kinase deficiency*. This is a key enzyme of the polyprenly pathway that biosynthesises polyprenly compounds, such as cholesterol, vitamin Ds and bile salts (*Figure 5.9*). The link between MK deficiency and hyper concentrations of IgD has not yet been established. Most sufferers are of western European origin, with many being Dutch or French. Clinical findings include, usually, increased concentrations of serum IgD, sometimes with higher levels of IgA. Patients suffer recurrent attacks of fever and symptoms of inflammation and an acute phase response from about the age of one year. These attacks may last up to six days and may be triggered by surgery, trauma or vaccinations. However, the severity of the phenotype is determined by residual levels of MK activity, giving a spectrum between these inflammatory symptoms and mevalonate aciduria. The latter is characterised by reduced psychomotor development and cerebellar ataxia and early death. Treatment using the recombinant interleukin (IL)-1 receptor antagonist (IL-1Ra) anakinra to counter the effects of the proinflammatory IL-1β has met with some success.

Hyper-IgM syndrome arises from mutations in *CD40* sited at 20q13.12 and in *CD40L*, which has been mapped to Xq26. The gene, *CD40*, codes for the CD40 protein expressed by B cells. Expression of *CD40L* in T cells produces the integral membrane protein CD40L, which is the ligand for CD40. Defects in *CD40L* are the more common. Interactions between the CD40L and CD40 are therefore prevented in hyper-IgM syndrome, but those between B cells and T-independent antigens can occur, leading to increased circulating concentrations of IgM. As may be expected, patients (children) suffer recurring viral and fungal infections.

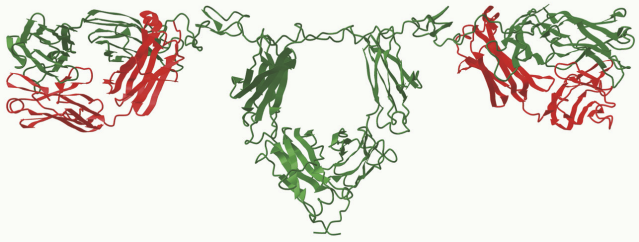


Figure 5.8 Molecular model of a solution (extended) structure of immunoglobulin D (IgD). The heavy and light chains are illustrated in green and red, respectively. PDB file 1ZVO.

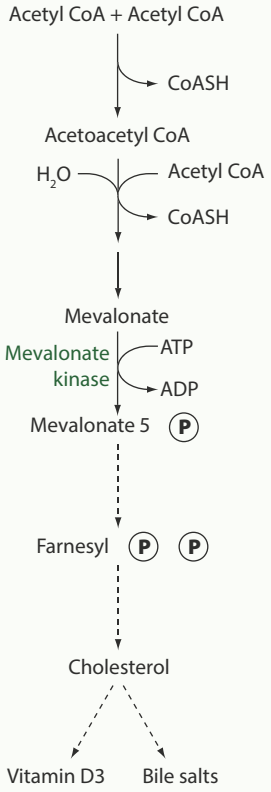


Figure 5.9 Simplified illustration of the polyprenly pathway, emphasising the reaction catalysed by mevalonate kinase. CoASH, coenzyme A (see *Chapter 10*).

It is essential that patients with XLA are diagnosed and treated as early as possible. Patients present with recurrent pyogenic infections, which occur generally from three to nine months of age, when maternally derived IgG concentrations become low. Infections encountered may result in pneumonia, otitis media, meningitis and diarrhoea and can be treated with antibiotics. The condition cannot be cured, but prognosis for children diagnosed before the age of five years is good using IRT. Tests for mutations in *BTK* are available, which allow a prenatal diagnosis using fetal cells obtained by chorionic villus sampling or amniocentesis (*Chapter 15*). Pregnant women can therefore receive genetic counselling and are offered the choice of a therapeutic abortion.

DiGeorge anomaly

DiGeorge anomaly (DGA), also referred to as DiGeorge syndrome or DiGeorge disease, is a developmental disorder involving organs that develop from the third and fourth pharyngeal pouches of the embryo during week 12 of pregnancy. Its incidence has been estimated to be between 1 in 4000 to 1 in 5000. DiGeorge anomaly is caused by deletion mutations of chromosome 22 (*Chapter 15*) at 22q11.2, which encodes over 60 genes. Most deletions occur *de novo*, with parents of children with DGA lacking genetic abnormalities.

Patients with DGA present with a broad range of variable phenotypes. These include congenital heart defects that may be so severe as to be life threatening, an absent or underdeveloped thymus resulting in cellular immune dysfunction of varying severities, parathyroid hypoplasia producing hypoparathyroidism and hypocalcaemia with symptoms of convulsions and tetany, and oral and facial abnormalities such as a cleft palate and abnormally widely spaced, downward slanting eyes.

Patients suffer severe and recurring viral and fungal infections. After heart conditions, immunological defects are the second most common cause of death in DGA patients. The number of circulating T lymphocytes is severely reduced, leading to defects in cell-mediated immunity. T cell proliferative responses to mitogens vary in DGA patients, such that they can be classified as either partial or incomplete. In the former, proliferation is reduced but, in the latter, it is completely absent. The absence of helper T lymphocytes reduces antibody production, so that antibacterial immunity may also be compromised, even though the number of circulating B lymphocytes is normal.

A diagnosis of DGA is based on reduced circulation numbers of T lymphocytes and their impaired proliferative response to mitogens, the cardiac malformations, hypoparathyroidism resulting in extended hypocalcaemia and a small or absent thymus. Fluorescence *in situ* hybridisation (FISH) detects deletions in chromosome 22 in most patients (*Margin Note 5.4*) but many small deletions are not apparent using standard FISH probes.

The prognosis for patients with DGA is variable and depends mostly on the degree of cardiovascular abnormalities and loss of thymic function. Treatment of DGA requires a range of specialisms. Cardiac malformations and oral defects must be rectified surgically. The hypocalcaemia is treated with calcium and vitamin D supplements. Children lacking any thymus function usually die by the age of two years due to the severity of their infections. Athymia can be addressed by T cell-replete HSCT (see *Box 5.2*) or a thymus transplant. The former can only result in post-thymic T cells but, nevertheless, has had some success, particularly when using a matched sibling transplant (*Chapter 6*). The surgical procedure of thymic transplantation is relatively straightforward. Results are, however, mixed: although some development of naive T cells that respond proliferatively to mitogens (see *Box 5.1*) is achieved in most cases, patients frequently experience autoimmune sequelae.

Wiskott–Aldrich syndrome, X-linked thrombocytopenia and X-linked congenital neutropenia

Wiskott–Aldrich syndrome, X-linked thrombocytopenia (XLT) and X-linked congenital neutropenia (XLN) all occur because of more than 300 mutations in *WAS*. They are often referred to as ***WAS*-related disorders**. The *WAS* gene is located on the short arm of the X chromosome at Xp11.22-p11.23. The gene is expressed only in nonerythroid haemopoietic cells, where it codes for the WAS protein (WASP). This protein is a key regulator in the polymerisation of cytoskeleton actin filaments in response to cellular signals.

Wiskott–Aldrich syndrome is a cytoskeletal (*Chapter 16*) disorder that affects blood platelets and lymphocytes. Its incidence is approximately 1 in 50 000–250 000. The pathogenesis of WAS is only partially understood but it is thought to arise from defective megakaryocyte functions and increased catabolism of platelets in the spleen (*Chapter 13*).

Margin Note 5.2
The 22q11 syndromes

DiGeorge anomaly is only one of several syndromes collectively termed 22q11 deletion syndrome. These include Cayler cardiofacial, conotruncal anomaly face, Sedlackova and Shprintzen–Goldberg syndromes. Their varying phenotypes, despite the common genetic aetiology, have led to varying nomenclatures and confused diagnosis of DGA.

Margin Note 5.3
Fetal alcohol syndrome

Other syndromes, with no apparent genetic link but with known environmental causes, bear some resemblance to DGA. One example is **fetal alcohol syndrome**, which results from prolonged exposure to alcohol during fetal development. Children with fetal alcohol syndrome also show the characteristic facial features associated with DGA.

Margin Note 5.4
Fluorescence *in situ* hybridisation (FISH)

Fluorescence *in situ* hybridisation (FISH) relies on the ability of fluorochrome-labelled DNA probes to hybridise with complementary DNA in tissue sections. The hybridised probe can be seen as fluorescent 'spots' in the nuclei of target interphase cells and can be located to specific chromosomal sites (*Figure 5.10*).

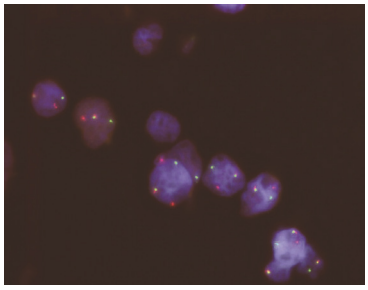


Figure 5.10 The use of FISH in diagnosing lymphomas. Red- and green-labelled probes are used that bind each side of an intact *BCL2* (*Chapter 17*) in a cell and would therefore give two yellow 'fusion' signals. Rearrangements of *BCL2* associated with the development of a lymphoma, as in the case here, means the probes are separated, resulting in a pattern of single red and green signals. Courtesy of Dr Nick Telford and Clare Hodgson, Oncology Cytogenetics, The Christie Pathology Partnership, The Christie NHS Foundation Trust, Manchester, UK.

Wiskott–Aldrich syndrome is characterised by a range of phenotypes. An absence of *WAS* expression results in the classical syndrome. Partial expression results in the milder XLT (*Chapter 13*). Both are characterised by low platelet numbers and volumes, which range from severe (≤20 000/mL; normal range 150 000–450 000/mL) to mild (≥50 000 mL); mean platelet volume (MPV) can be as low as 5 fL, compared with the normal range of 8.9–11.8 fL. X-linked congenital neutropenia is caused by gain-of-function mutations in *WAS*. It is characterised by low numbers of circulating neutrophils, with patients suffering recurrent life-threatening infections.

Classical WAS is associated with decreased concentrations of IgM but often with increased production of IgE and IgA. Since IgM is the prevalent antibody in immune responses to bacterial polysaccharides, there is an increased incidence of infections with encapsulated bacteria. The thymus is normal at birth but subsequently decreases in size. Thus, in the early stages, T and B cell numbers are normal, but, as WAS progresses, there is a loss of both humoral and cell-mediated immunities. There are increased incidences of leukaemias, lymphoid tumours and autoimmune disorders with age.

Patients with WAS present in their first year with recurrent infections and low platelet numbers that typically result in thrombocytopenia, leading to increased bleeding times (*Chapter 13*) that can prove fatal, petechiae, a tendency to bruise easily, and eczema. The use of full blood analyses to aid diagnosis can be inconclusive given the range of phenotypes presented by WAS (and XLT). The use of flow cytometry (*Chapter 6*) analysis of *WAS* expression is a recommended diagnostic method.

Treatment of WAS involves IRT and antibiotics to provide protection against infections. Platelet transfusions to prevent bleeding have been successful in some cases. Bone marrow transplants or HSCT (see *Box 5.2* and *Chapter 6*) are the only curative therapies, but gene therapy trials are currently under investigation. Unfortunately, without an HSCT (see *Box 5.2* and *Chapter 6*), the prognosis for WAS patients is poor, with death commonly occurring before the age of four years, usually from severe infections and bleeding. However, a successful BMT/HSCT is associated with long-term survival. X-linked thrombocytopenia has similar bleeding risks to WAS but is not associated with the other significant clinical features. It can be treated more conservatively with antibiotics, antibleeding therapies, IRT and topical steroids for the eczema. X-linked congenital neutropenia is treated with recombinant GCSF (*Figure 5.5*) to stimulate increases in the numbers of neutrophils in the blood.

Genetic counselling is recommended for women who have had a child with a *WAS*-related disorder. Detection of mutated *WAS* in cells obtained by chorionic villus sampling or amniocentesis in future pregnancies allows a prenatal diagnosis, with the possibility of terminating the pregnancy if the fetus is found to be affected.

Chédiak–Higashi syndrome

Chédiak–Higashi syndrome (CHS) affects the lysosomal transport and storage functions of phagocytic cells (see *Chapter 16*). It is a recessive autosomal disorder first described in 1943; even compared with many IEIs it is a rare disorder, with only several hundred cases reported worldwide. The syndrome arises from mutations in *LYST*, which is located at 1q42.1-q42.2. This gene encodes the lysosomal trafficking regulator protein (LYST) that is essential for the transport of proteins into lysosomes and for lysosomal maturation. Mutations result in neutropenia and abnormal neutrophils containing large, densely staining granules but which fail to kill ingested bacteria, and platelets with abnormally dense bodies. For this reason, it can be regarded as a phagocytic defect. Natural killer cells (*Chapter 4*), which form the first line of defence against viruses, are also abnormal. In addition, melanocytes, the pigment-producing cells of the skin, contain larger than normal melanosomes (the pigment-storing organelles), resulting in abnormal deposits of the pigment (*Figure 5.11*).

Patients with CHS suffer recurrent bacterial infections, most commonly with *S. aureus*, *Streptococcus pyogenes* and *Pneumococcus* spp., and viral infections, with, for example, Epstein–Barr virus, that frequently result in lymphomas. Infants suffer recurrent skin infections, which may result in ulcers and abscesses. The abnormalities of melanocytes mean that patients are deficient in skin pigment and have blond hair, often with a metallic lustre, and translucent blue eyes (oculocutaneous albinism) with impaired vision. The defective platelets mean that infants bruise easily. In addition, they suffer progressive neurological dysfunction, with abnormal gait, intellectual disabilities and peripheral neuropathy. If the child survives beyond the first decade this may lead to Parkinsonism and/or dementia.

Children exhibiting immune deficiencies, congenital neutropenia and abnormal pigmentation of hair, skin and eyes should be investigated for CHS. The presence of abnormally large cytoplasmic granules in relevant blood cells is diagnostic *(Figure 5.11)*. Molecular techniques can also identify mutations in *LYST*.

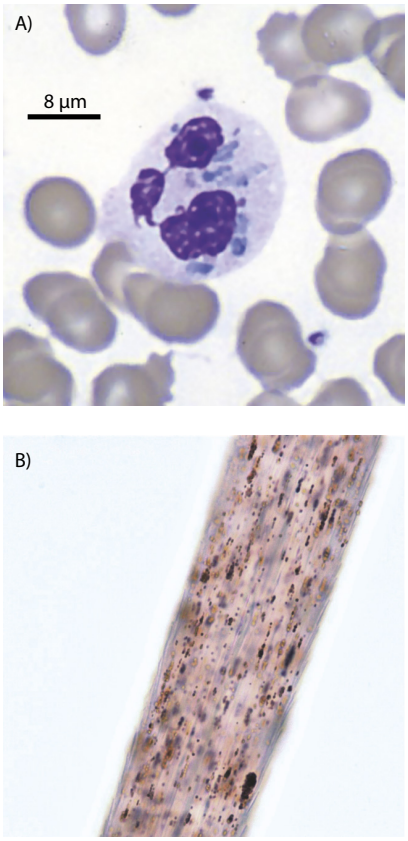


Figure 5.11 Photomicrographs showing (A) the large densely staining granules in a neutrophil and (B) a hair shaft with abnormal clusters of clumped melanin granules. Both photographs are from Chédiak–Higashi syndrome patients. Courtesy of Glenn Anderson, Great Ormond Street Hospital for Children, London, UK. ⏎

Morbidity and mortality are high in CHS, with untreated children frequently dying before the age of 10 years, usually from overwhelming pyogenic infections attributable to poor neutrophil function. Antibiotics are given to treat bacterial infections and antiviral drugs (*Chapter 3*) to limit infection with the Epstein–Barr virus, and interferon-α (IFN-α) to try and restore some immune function. The recommended treatment for children with CHS is a HSCT (see *Box 5.2* and *Chapter 6*) to correct the immune deficit (see *Box 5.2* and *Chapter 6*). However, even when this treatment is successful and immune functions are promoted, it does not correct neurological conditions or the lack of pigmentation, since these defects arise from cells not derived from bone marrow.

Given its autosomal recessive cause, both parents must be heterozygous for the mutated gene. Therefore, genetic counselling is recommended for families who have had a child with CHS. Prenatal diagnosis allows the possibility of terminating future pregnancies if the fetus is affected.

Phagocytic defects

Chediak–Higashi syndrome emphasises the importance of correctly functioning phagocytic cells, monocytes, macrophages and neutrophils, which kill ingested bacteria using several different mechanisms (*Chapter 4*). Defects in any of these mechanisms increase the incidence of infections, even when B and T cell populations and functions are normal. Examples of such defective phagocytes are seen in chronic granulomatous disease (CGD) and leukocyte adhesion deficiencies (LADs).

Chronic granulomatous disease is named from the granulomatous inflammatory nodules that present on the skin and in the gastrointestinal and genitourinary tracts. Its incidence is estimated to be 1 in 200 000–250 000. About 70% of cases are X-linked; the remainder are inherited in a recessive autosomal fashion.

Patients have defective NADPH oxidase activity that impairs their ability to generate the reactive oxygen intermediates, such as superoxide radicals and hydrogen peroxide (*Figure 5.12*), needed to kill bacteria ingested into phagosomes. Reduced nicotinamide adenine dinucleotide phosphate (NADPH) oxidase is a multimeric enzyme: defects in any subunit can result in CGD. The cytochrome b_{558} (also called b_{-245}) subunit is encoded by *CYBB* located on the X chromosome; genes for other subunits are located on autosomal chromosomes, which accounts for the patterns of inheritance.

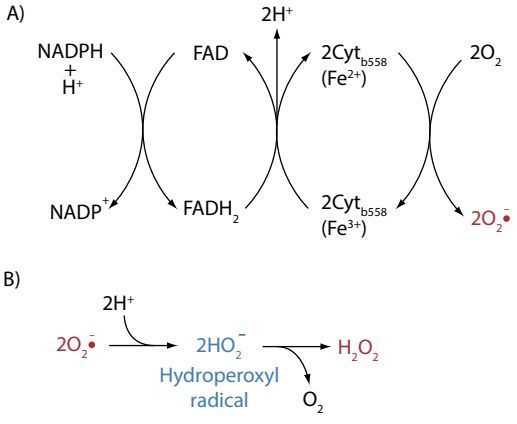


Figure 5.12 (A) The use of NADPH in the formation of superoxide radicals. (B) The conversion of superoxide radicals to hydrogen peroxide. Cyt_{b558}, cytochrome b_{558}.

Children with CGD usually present before the age of three years. Skin infections, pneumonia, gastroenteritis and perianal abscesses are common. Pneumonia is generally associated with fungal infections; disseminated fungal disease also occurs. Common causes of infections include *Burkholderia cepacia*, *Nocardia* spp., *Serratia marcescens* and *S. aureus*. *Aspergillus* spp. are the most common causes of fungal infections. Chronic inflammation results in the formation of granulomas, which are discrete granules composed of several types of accumulated cells. Abscesses on the lungs, gastrointestinal tract, spleen and liver may also be present.

A diagnosis of CGD considers the early onset of recurrent infections, the development of granulomas, hepatosplenomegaly (that is, enlarged liver and spleen) and lymphadenopathy. Laboratory investigations include the nitroblue tetrazolium (NBT) test to determine the activity of the NADPH oxidase. In neutrophils with normal levels of enzyme, the pale yellow-coloured NBT is reduced to a blue compound as NADPH is oxidised and this can be observed in the cytoplasm by microscopy. Flow cytometry methods are also employed to assess activity.

Patients with CGD are treated with large doses of antibiotics and antifungal compounds over long periods of time (*Chapter 3*). This treatment also helps to dispel the granulomas. Treatment with IFN-γ induces tumour necrosis factor-α (TNF-α) production, which inhibits inflammation and the formation of granulomas. Abscesses may need to be drained. However, the only cure is a successful HSCT (see *Box 5.2* and *Chapter 6*).

Leukocyte adhesion deficiencies result from the failures of leukocytes to express cell adhesion molecules essential for chemotaxis and movement through blood vessel walls to sites of inflammation. Phagocytes are therefore unable to enter inflamed tissues to ingest bacteria.

Three LADs, numbered I, II and III in order of their chronological 'official' descriptions, have been described. All three result from autosomal recessive mutations, but the genes involved differ in each type. Mutations in *ITGB2*, located at 21q22.3, mean that neutrophils, macrophages and lymphocytes do not express β2 integrin (CD18), a leukocyte cell adhesion molecule (*Margin Note 5.5*) and results in LAD I. Integrins are an evolutionary old family of cell adhesion molecules. All are αβ heterodimers. Both monomers are integral membrane glycoproteins, with large extracellular domains. Eighteen different α and eight β subunits have been described, but only 24 different combinations are known in humans (*Figure 5.13*). Integrins are receptors that mediate the attachment of cells to the extracellular matrix (*Chapter 16*), such as when neutrophils bind to endothelial cells. The β_2 monomer is only expressed on leukocytes. It forms heterodimers with four different types of α subunit allowing interactions with different proteins and binding to endothelial cells lining the blood vessels. In addition, CD18 is a receptor for C3b, an opsonin for phagocytic cells in complement activities. Leukocyte adhesion deficiency II is caused by mutations in *SLC35C1* at chromosome location 11p11.2. This gene encodes GDP-fucose transporter 1, which moves GDP-fucose into the lumen of the endoplasmic reticulum and Golgi apparatus (*Chapter 16*). The defects in fucose metabolism in LAD II mean that the biosynthesis of the CD15 antigen (sialyl Lewis X glycan; *Figure 5.14* and *Chapter 6*) is impaired. This antigen is required by leukocytes to facilitate their binding to E and P selectins (*Margin Note 5.5*) on endothelial cell surfaces: a prerequisite for them to be able to 'roll along' the endothelium and travel to sites of infection and inflammation.

Margin Note 5.5 Cell adhesion molecules

Integrins are only one of four families of cell adhesion molecules:

- cadherins
- immunoglobulin-like adhesion molecules
- integrins
- selectins

Figure 5.13 Molecular model of the extracellular domain of an αβ integrin dimer. The α and β subunits are shown in purple and light purple, respectively. The red-brown spheres represent bound Ca^{2+}. The approximate position of the plasma membrane is shown in grey. PDB file 3IJE.

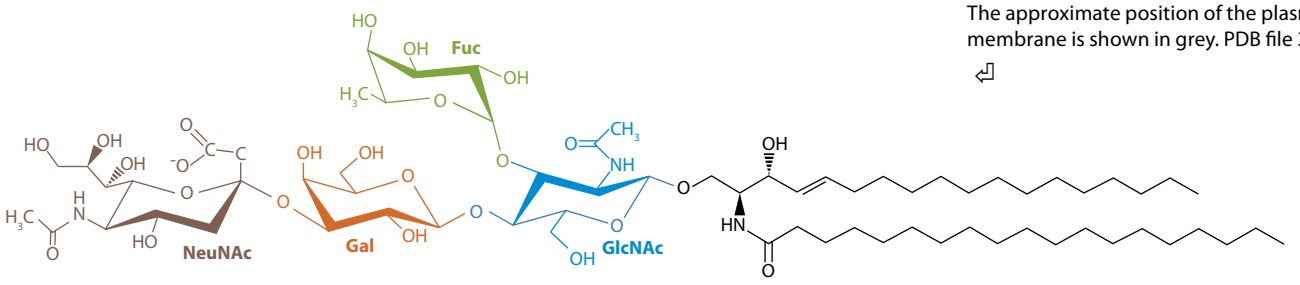


Figure 5.14 A sphingolipid carrying the CD15/sialyl Lewis X tetrasaccharide/glycan.

Leukocyte adhesion deficiency III occurs following mutations in *FERMT3* at the cytogenetic location 11q13.1. This gene is specifically expressed in haematopoietic cells where it codes for kindlin 3, a protein involved in the interactions between integrins and the cytoskeleton.

Symptoms of LAD I vary in severity between individuals. Patients usually suffer a delayed detachment of the umbilical cord (at approximately four or five, rather than two, weeks) and bacterial and fungal infections, particularly of the skin and mucous membranes, that may become life threatening. Pus is absent at sites of infection. Untreated infants with LAD I do not usually survive beyond two years of age. Patients with LAD II usually develop recurrent bacterial infections, although these are usually less severe than in LAD I. Indeed, in general, the frequency of infections decreases from the age of three years. Symptoms of LAD III resemble those of I, but unfortunately patients also suffer bleeding problems, such as excessive nose bleeds and large red/purple spots on the skin arising from subcutaneous bleeding, from birth.

Diagnosis of suspected LADs may be confirmed by flow cytometry, to assess the presence of the cell adhesion molecules on blood leukocytes, or by molecular biological techniques to identify the specific mutations. Recurrent infections require antibiotics and antifungal treatments, but the only cure is a HSCT (see *Box 5.2* and *Chapter 6*). Prenatal diagnosis is also available to allow for a therapeutic abortion. Affected families require genetic counselling.

Complement deficiencies

A deficiency of any complement protein leads to increased susceptibility to infections and can severely compromise health. Defects are relatively common, estimated at 1 in 1000, constituting up to 10% of all IEIs. The majority are inherited as autosomal mutations of numerous different genes, although X-linked patterns have also been reported for mannose-binding lectin (MBL) and properdin.

Complement is also involved in the pathology of several immunological disorders, including autoimmune diseases, such as rheumatoid arthritis and autoimmune haemolytic anaemias (*Section 5.4*).

Inherited deficiencies of all complement proteins have been described. Given the role of complement in removing immune complexes from the circulation, deficiencies in several early proteins of the classical pathway result in increased incidences of immune complex disorders. Patients also suffer recurring infections, particularly with staphylococci and streptococci. Over 90% of patients with a homozygous C1q deficiency and 10% of patients with a homozygous C2 deficiency develop SLE (*Section 5.4*).

A deficiency of MBL (*Figure 5.15*), the most common of the complement associated IEIs, is associated with increased incidence of infections with, for example, *Neisseria* spp., *Pseudomonas aeruginosa* and the yeast *Saccharomyces cerevisiae.* Newborn babies and infants are particularly at risk of this deficiency, indicating the importance of this pathway in protecting the young from infection. Deficiencies in alternative pathway proteins, such as complement factors B and D and properdin, lead to defective opsonisation and recurrent *Neisseria* infections. Patients with defective terminal complement pathway proteins usually present clinically at an older age than sufferers of other complement defects.

Complement is potentially capable of clinically damaging inflammatory actions. Fortunately, complement proteins are extremely labile and have extremely short half-lives. Also, numerous soluble and membrane proteins control its actions, for example the C1 inhibitor (C1INH), which prevents the unregulated activation of complement (*Box 5.4*), and decay-accelerating factor (DAF), which prevents damage to bystander host cells. Suspected complement deficiencies in the classical and the alternative pathways can be investigated using the CH_{50} and AP_{50} tests, respectively. The former measures the ability of patient serum to lyse antibody-coated sheep erythrocytes and the latter its ability to lyse uncoated cells in rabbits. Results are expressed in terms of the ability of the serum to lyse 50% of the target erythrocytes. Individual complement protein concentrations can be determined by immunoassay, using specific antibodies. Thus, C3 and C4 can be measured by nephelometry, single radial immunodiffusion or enzyme-linked immunosorbent assay (ELISA; *Chapter 4*). For example, healthy individuals have MBL blood concentrations of approximately 1 μg/mL (which can increase up to fourfold during inflammation), while a deficiency is defined as less than 0.5 μg/mL.

Treatment involves antibiotics to manage infections. All routine vaccines should be administered. The use of specific targeted therapies with monoclonal antibodies to modulate complement activities is currently receiving attention. Counselling to make patients aware of the symptoms of meningococcal infection and the need to seek immediate help should they develop is essential.

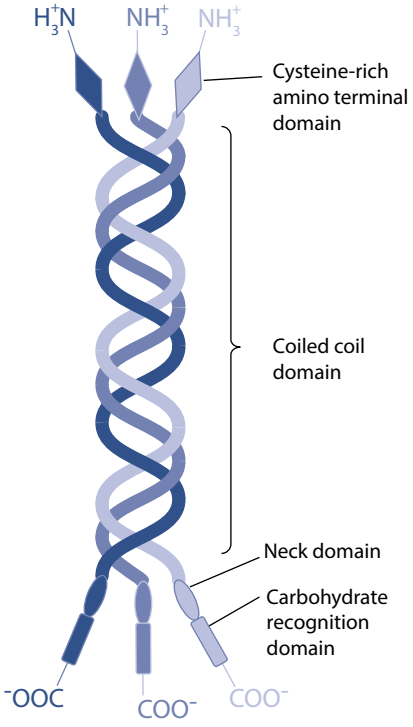


Figure 5.15 Cartoon illustrating the four domains of a mannose-binding lectin (MBL) molecule. ⏎

BOX 5.4 HEREDITARY ANGIOEDEMA ⏎

Hereditary angioedema (HAE) is characterised by episodic oedemas that can involve any part of the body. The condition has an incidence of 1 in 50 000. Two types of HAE arise from autosomal dominant mutations in *SERPING1* at 11q12.1. This gene encodes for the complement regulator C1 inhibitor (C1INH), which normally functions in preventing the overactivation of the first part of the classical pathway by inhibiting C1r and C1s. A lack of inhibitor leads to the consumption of C4 and C2. Type I constitutes 85% of cases and results from reduced concentrations of C1INH, while in type II the inhibitor is present but nonfunctional. A third form, HAE III, arises from rare mutations in *F12* at 5q35.3, which codes for clotting factor XII (*Chapter 13*). Clinical episodes in HAE III are indistinguishable from those of I and II. Like HAE I and II, mutations causing type III are inherited in an autosomal dominant fashion but the concentration and function of C1NH are normal, as are those of C4.

The disorder presents as noninflammatory swellings of the skin, especially that of the limbs, which is often precipitated by physical trauma and anxiety. Abdominal pain is caused by the involvement of internal organs, such as the stomach, intestines and bladder. Severe oedema of the larynx or obstruction of the airways can cause death. Treatment during attacks is to administer fresh frozen plasma or commercially available C1INH. Prophylactic treatment with danazol, an androgen, has been shown to stimulate the synthesis of C1INH, but prolonged treatment may result in unpleasant side effects, such as virilisation in women and suppression of testosterone production in men (*Chapter 7*).

5.4 AUTOIMMUNE DISORDERS

Autoimmune diseases result from a dysfunctional immune system that produces antibodies directed against self antigens on the patient's own proteins and cells. Clearly, their development arises from failures of immunological tolerance (*Box 5.5*). Autoimmune diseases affect 5–7% of the population. Many autoimmune disorders affect about two to four times as many female as male patients. The reasons for this gender bias are unclear but may be related to concentrations of sex hormone (*Chapter 7*). The risk of developing one of these conditions is, at least partly, genetic, with some conditions running in families. Many autoimmune disorders show a link with the type of histocompatibility complex (HLA; see *Chapters 4* and *6*) antigens present on cells. For example, patients with Goodpasture syndrome have a higher incidence of HLA-DR2 than the healthy population.

Generalised symptoms associated with autoimmune diseases include tiredness, skin rashes, pains in joints and abdomen, swollen lymph glands and recurrent fevers. They can affect virtually any part of the body and are painful, chronic and debilitating.

Autoimmune disorders are often classified according to whether they are organ-specific or systemic (*Table 5.5*). Like most classifications, especially in biology(!), that of autoimmune disorders is imperfect. For example, Goodpasture syndrome affects both kidneys and lungs while multiple sclerosis (MS) exerts systemic effects but by attacking only one type of tissue.

Destruction of cells and tissues can be effected by autoantibodies and/or cell-mediated immunity. For example, in MS, patients produce antibodies against the myelin surrounding axons, but they also have T_H and T_C lymphocytes in their blood and cerebrospinal fluid specific for myelin protein. Thus, humoral and cell-mediated autoimmunity may contribute to the demyelination of nerves in MS patients. In other conditions, autoantibodies can block or stimulate a cell receptor, such as in myasthenia gravis and Graves' disease, respectively (see below and *Chapter 7*).

Autoimmune diseases are a heterogenous group of disorders and often difficult to diagnose. More than 100 types have been reported, so, as with IEIs, only a restricted number will be highlighted to illustrate their range.

BOX 5.5 HOW TO RECOGNISE SELF ⏎

Several theories have been proposed to explain the development of immunological tolerance and, indeed, there are likely to be several different mechanisms involved. In experimental systems, rodents become tolerant to potential immunogens if they are exposed to them during fetal development. Thus, mice exposed to foreign proteins *in utero* do not mount immune responses against these immunogens when adults. This suggests that lymphocytes exposed to epitopes, including self epitopes, during fetal development are selectively removed or deleted from the immune system, which explains the experimental induction of tolerance. However, there is certainly evidence that this happens to T lymphocytes developing in the thymus in human fetuses (*Figure 5.16*). It does not explain the development of tolerance to immunogens that are not present in the fetus but which are expressed in the adult. It seems, at least in some cases, that some immunogens are kept anatomically separate from the immune system during life, to avoid provoking an autoimmune response. For instance, the anterior chamber of the eye and lens are considered sequestered from the lymphatic system. However, when rabbits are injected with lens protein, they make antibodies that then bind to the lenses of their own eyes. Another example is seen in vasectomised men, who may start to make antispermatozoa antibodies, presumably because the spermatozoa they continue to produce become exposed to their immune systems following the operation. Finally, there is evidence that some types of T lymphocytes can suppress immune responses against self antigens. These T lymphocytes have, in the past, been called suppressor T lymphocytes and were thought to belong to the CD8+ subset. However, it has been shown that CD4+ and CD8+ cells can have suppressor activity by producing inhibitory cytokines, such as interferon-γ and interleukin-10, respectively.

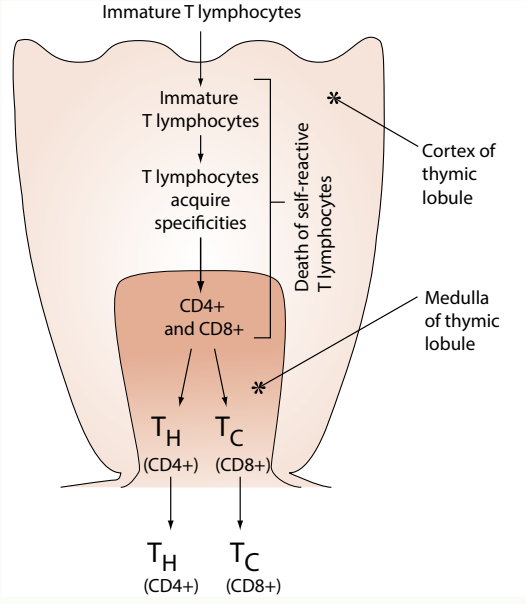


Figure 5.16 Schematic showing the development of T lymphocytes. Immature T lymphocytes enter the thymic lobule and mature into helper and cytotoxic subsets as they progress through the thymic lobules. In the cortex, cells that recognise self epitopes are deleted. ⏎

Type of disorder	Example	Effect of disorder	Autoantibodies present	Autoreactive T cells
Organ specific	Autoimmune haemolytic anaemia	Destruction of erythrocytes (*Chapter 13*)	Antibodies to erythrocyte antigens	
	Autoimmune or Hashimoto thyroiditis	Hypothyroidism (*Chapter 7*)	Antibodies to thyroglobulin and thyroid peroxidase	T_H1 cells specific for thyroid antigens
	Addison disease	Adrenal insufficiency (*Chapter 7*)	Antibodies to cytoplasmic antigens of cells of adrenal cortex	Infiltration of adrenal cortex with autoreactive T cells
	Type 1 diabetes mellitus	Destruction of insulin-producing cells in pancreas; serious metabolic disturbances	Antibodies to islet cells found in classical juvenile form	Infiltration of pancreas with autoreactive T cells
	Goodpasture syndrome	Progressive kidney and lung damage	Antibodies to basement membrane antigens of kidney and lung	
	Graves disease	Hyperthyroidism (*Chapter 7*)	Antibodies to thyroid-stimulating hormone receptors	Destruction of thyroid cells by autoreactive T lymphocytes
	Myasthenia gravis	Progressive muscle weakness	Antibodies to acetylcholine receptors on muscle cells	
	Pernicious anaemia	Failure to absorb vitamin B_{12} in the stomach (*Chapters 11* and *13*)	Antibodies to intrinsic factor	
Systemic disease	Rheumatoid arthritis (RA)	Inflammatory disorder affecting joints, skin and internal organs	Antibodies to immunoglobulin G (IgG) (rheumatoid factor)	Infiltration of joints with autoreactive T lymphocytes
	Systemic lupus erythematosus (SLE)	Inflammatory disorder affecting multiple organ systems	Antibodies to DNA, chromatin and histones; rheumatoid factor in some individuals	Evidence of T cell reactivity in some of the many organs affected
	Multiple sclerosis (MS)	Inflammatory disorder affecting central nervous system	Antibodies to myelin basic protein	Destruction of myelin membrane by autoreactive T lymphocytes

Table 5.5 Examples of autoimmune diseases. ⏎

Autoimmune disorders affecting endocrine glands

Autoimmune disorders of the thyroid gland are among the most common autoimmune disorders. Some will be discussed in *Chapter 7*.

Autoimmune thyroiditis or Hashimoto thyroiditis

Autoimmune thyroiditis or Hashimoto thyroiditis is an infiltration of CD4+ and CD8+ T lymphocytes that progressively destroy the thyroid gland, resulting in hypothyroidism and myxoedema. Its cause is not fully understood but it has been linked with several genetic lesions. The condition presents typically in middle-aged women who are overweight, lethargic, constantly feel cold and constipated and have coarse, dry hair and skin. The thyroid is swollen with a generally painless goitre, but which has a rubbery consistency when palpated. Patients have antibodies to thyroglobulin and to thyroid peroxidase, hence the decreased concentration of thyroxine. Diagnosis can be confirmed by a biopsy of the gland and can also be investigated by ELISA. Immunohistochemical techniques (*Chapter 4*) can show the patient to have antibodies that bind to microsomal antigens in sections of normal thyroid tissue (*Figure 5.17*). Patients can be treated with thyroxine replacement therapy. Thyroidectomy may be required. The prognosis for patients with Hashimoto thyroiditis is good.

Graves' disease

Graves' disease is a multisystem autoimmune disorder characterised by hyperthyroidism. Its cause is poorly understood but has a genetic basis. Patients have nodules in the thyroid that are foci of infiltrating T lymphocytes. More than 90% of patients have antibodies to thyroid-stimulating hormone (TSH) receptors on the surface of thyroid cells.

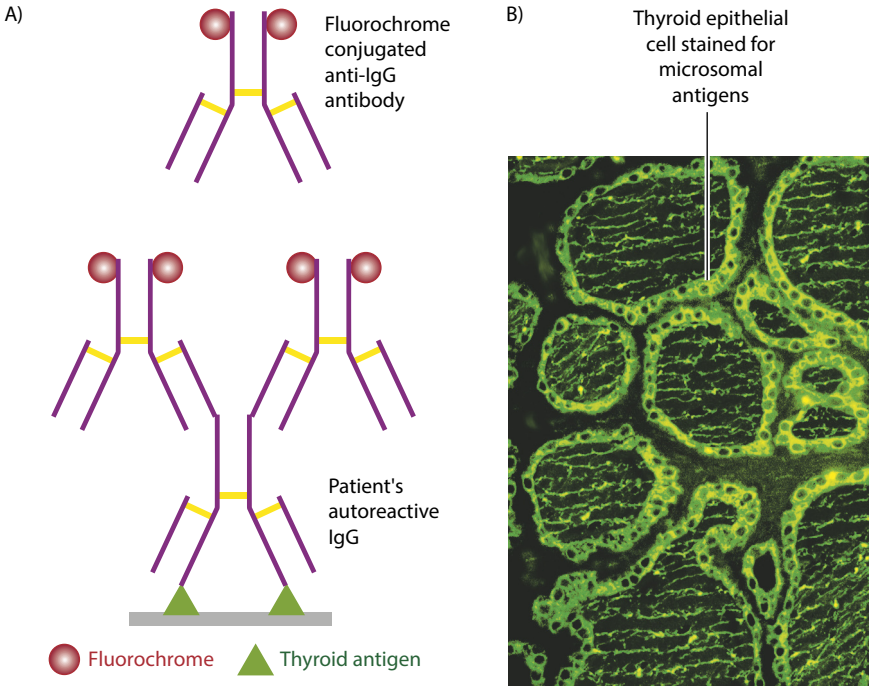


Figure 5.17 Immunofluorescence in autoimmune thyroiditis. (A) Schematic showing the basis of indirect immunofluorescence staining of thyroid antigens by a patient's serum. IgG, immunoglobulin G. (B) Photomicrograph showing thyroid microsomal antigens stained with a patient's serum by indirect immunofluorescence. Courtesy of EUROIMMUN AG, Germany.

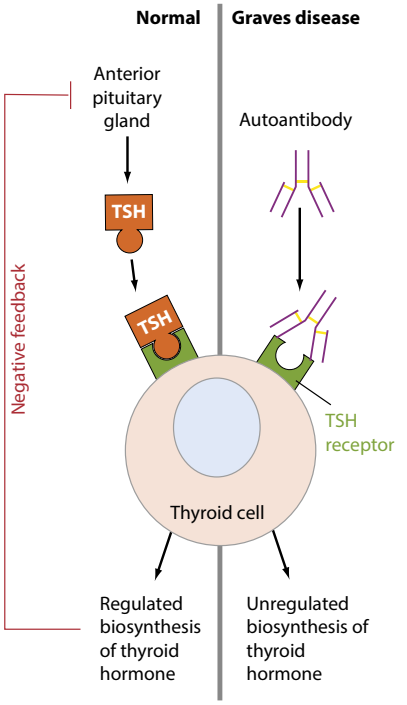


Figure 5.18 The left-hand side of the schematic shows how the production of thyroid hormone is normally regulated by a feedback mechanism; the right illustrates how the binding of an autoantibody to the thyroid-stimulating hormone (TSH) receptor interferes with this mechanism and can lead to Graves disease.

Binding of antibodies to receptors stimulates the overproduction of thyroid hormones (*Figure 5.18*) because it is not regulated by the usual negative feedback mechanisms (*Chapter 7*). Patients present with a pronounced thyrotoxicosis and bulging eyes (**exophthalmos**) due to the growth of tissue around the orbits, are thin, have high resting pulse rates, constantly feel hot and may suffer diarrhoea and general agitation. Low levels of antibodies to thyroid microsomal antigens are seen in the plasma.

Graves' disease can be diagnosed from the general symptoms presented and a biopsy of the thyroid tissue and confirmed by a thyroid scan that shows an uptake of radioiodine above reference values. Treatment may involve antithyroid medication, such as methimazole or propylthiouracil, to reduce the biosynthesis of thyroid hormones, and β-blockers to treat the tachycardia (*Chapters 13* and *14*). Destruction of thyroid tissue can be achieved by administering radioactive iodine that becomes concentrated in the thyroid or by surgery. Prognosis is usually excellent.

In pregnant women with Graves' disease, autoantibodies to TSH cross the placenta and consequently the baby is born with thyrotoxicosis. Urgent treatment is required but, in time, the baby recovers as the concentration of maternally derived antibodies decreases.

Insulin-dependent diabetes mellitus type 1 and Addison disease

Both insulin-dependent diabetes mellitus type 1 and Addison disease are discussed extensively in *Chapter 7*. The former is caused by an autoimmune destruction of the insulin-producing cells in the pancreas and the latter by autoimmune damage to the adrenal cortex. Both diseases are fatal unless treated by replacing the missing hormones.

Antiglomerular basement membrane disease

Antiglomerular basement membrane disease or Goodpasture syndrome is characterised by the production of autoantibodies to basement membranes (*Chapter 9*), particularly in those of the glomeruli (anti-GBM) and lungs. Its incidence is rare: about 0.5 cases per million in the UK. Unlike most autoimmune disease, it occurs predominantly in males. Tissue damage is caused by antibodies binding to basement membrane and activating a complement-mediated inflammation. This leads to glomerulonephritis causing proteinuria and haematuria, and erythrocyte casts (*Margin Note 5.6*). Without treatment, the condition leads to progressive renal failure. Damage to alveolar membranes causes **haemoptysis**, that is, the coughing up of blood from the lungs. About 40% of patients experience chest pain. Haemorrhaging from the lungs may eventually lead to respiratory failure. There is some suggestion that the binding of antibodies to the alveolar basement

Margin Note 5.6
Erythrocyte casts

Erythrocyte casts are formed in the kidneys when erythrocytes become trapped in hyaline and form deposits in the kidney tubules and are lost in the urine. Healthy individuals will have a few, consisting of hyaline in their urine, but finding casts containing erythrocytes or leukocytes is indicative of kidney disease.

membranes is facilitated by exposure to organic solvents, which increase the permeability of the alveolar capillaries.

Antiglomerular basement membrane disease can be confirmed in suspected patients from anti-GBM antibodies in serum samples or by ELISA methods on a kidney biopsy. An early diagnosis is essential and treatment must start immediately. Therapy involves removal of circulating antibodies by plasmapheresis (*Margin Note 6.6*) and the administration of immunosuppressive drugs (*Chapter 6*). The mortality for anti-GBM disease is improving, and usually survival rates correlate with the extent of renal function.

Myasthenia gravis

Myasthenia gravis (MG) is a chronic progressive neuromuscular disorder in which patients produce antibodies to acetylcholine receptors at the neuromuscular junctions of striated muscle. Blockage by the antibodies means the receptors fail to respond to acetylcholine (*Figure 5.19*). The incidence of MG has been quoted as about 14 in 100 000, with a female to male ratio of roughly 3:2. The condition results in an intermittent but progressive weakness of skeletal muscles, including those necessary for breathing and those involved in chewing, swallowing, talking and eye movements. Difficulty with respiration can lead to inadequate intake of oxygen and an inability to clear secretions from the respiratory tract with an increased incidence of pneumonia. Defective eye movements can lead to double vision and an inability to raise the eyelids, a condition known as **ptosis**. Approximately 75% of patients with MG also have thymic abnormalities such as **hyperplasia** (increased numbers of cells as a result of increased division) and thymoma (a, usually, benign tumour).

Problems with eye movements are often the first sign of MG. Diagnosis is aided by assessing any improvements in muscle strength following the intravenous injection of short-acting acetylcholinesterase inhibitors. The presence of autoantibodies may be confirmed by indirect immunofluorescence tests, and their concentrations are a useful measure of disease progression. However, autoantibodies may not be detected in patients when the disease is confined to the facial muscles. Patients are treated with immunosuppressive drugs and cholinesterase inhibitors, and by plasmapheresis to remove the autoantibodies. The condition is improved in the majority of patients by thymectomy. The mortality rate of MG patients is significantly improving.

Rheumatoid arthritis

Rheumatoid arthritis (RA) is a chronic, painful and debilitating inflammatory condition characterised by symmetrical arthritis and radiological changes to the bone. The aetiology of RA remains unknown, despite numerous infectious agents having been implicated over the years. It is, however, one of the most common autoimmune disorders, with an incidence of 1–2 per 100. The female to male ratio is approximately 3:1, and the condition generally appears at the of age of 40–60 years, although juvenile forms also occur.

Autoantibodies are present in the plasma of RA patients: the most common, occurring in 70% of cases, is an IgM class antibody called **rheumatoid factor** (RF) directed against IgGs. However, RFs belonging to IgA and IgG classes have also been detected. Approximately 40% of patients with RA also have **antinuclear factors**, which are also seen in sufferers of SLE (see below). The presence of RF causes large amounts of immune complexes to be formed, since IgG is present at relatively high concentrations in the blood. The complexes may adhere to blood vessel walls, activating complement and initiating inflammation. Patients may also suffer inflammation of blood vessels or vasculitis, and about 20% have subcutaneous rheumatoid nodules, often on the elbows and forearms, but which may also occur in internal organs. The nodules consist of a mass of monocytes, lymphocytes and plasma cells surrounding a necrotic core, and probably represent the progression of vasculitis.

While a minority of patients suffer a single episode of joint inflammation with long-term remission, most have a progressive illness characterised by alternating periods of remission and exacerbation ('flares'). In periods of active inflammation, the affected joints (*Figure 5.20*) are painful, swollen, red and warm to the touch, all symptomatic of the internal inflammation. The presence of RF and subsequent inflammatory disease does not, however, fully explain the pathogenesis of RA. Cell-mediated immunity is known to be heavily involved in joint destruction. Small lymphocytes, especially T_H1 cells, monocytes and macrophages infiltrate synovial membranes, leading to thickening of the membranes and swelling of the joints. Activated macrophages within the synovial fluid produce cytokines, such as IL-1 and TNF-α, which mediate erosion of bone. The accumulation of inflammatory neutrophils within the synovial fluid also contributes to damaging the cartilage.

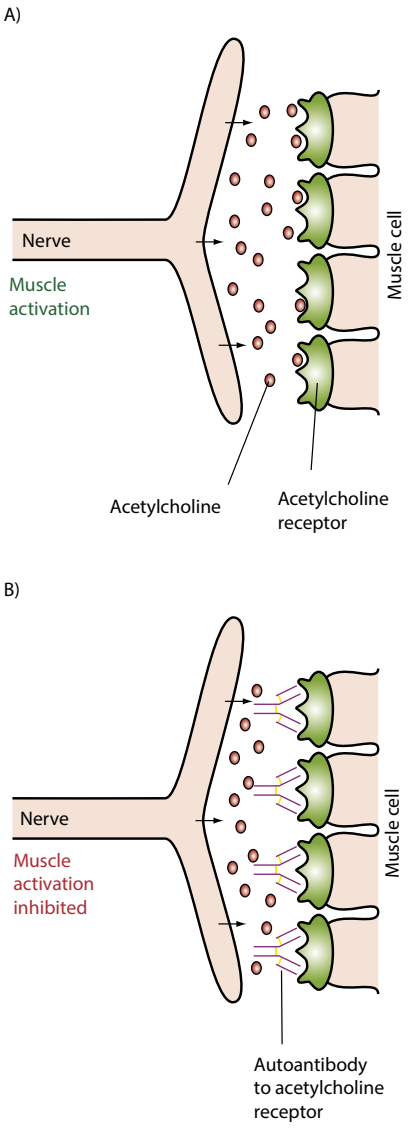


Figure 5.19 Schematics showing (A) the normal stimulation of muscle cells by the binding of acetylcholine to receptors on muscle cells and (B) the blocking of acetylcholine receptors by autoantibodies in myasthenia gravis.

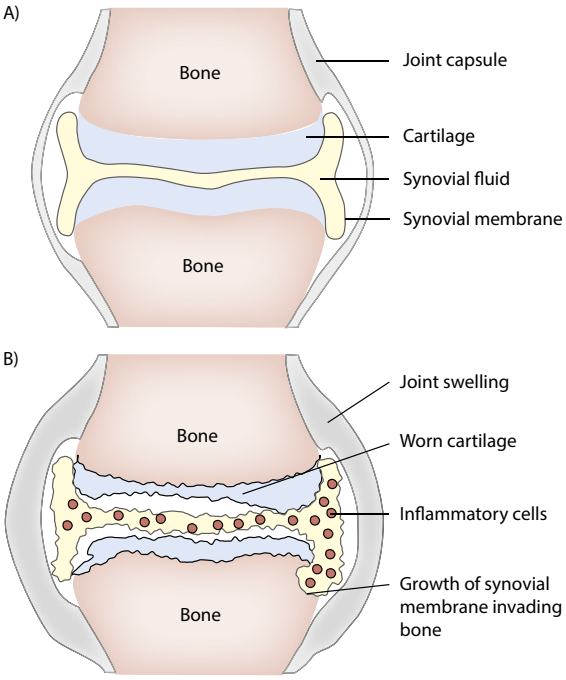
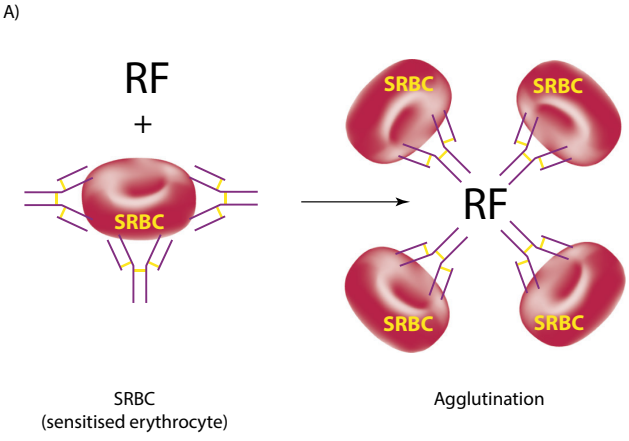


Figure 5.20 Schematics showing (A) a normal synovial joint and (B) the characteristic changes associated with rheumatoid arthritis. (C) The arthritic hands of an elderly person. Courtesy of Shutterstock.com/Chaowalit Seeneeha.

A diagnosis of RA must be achieved as early as possible so treatment can begin to mitigate long-term damage. Diagnosis requires an evaluation of numerous signs and symptoms, for example:

- stiffness of the joints in the morning
- arthritis in three or more joints
- arthritis of the joints in the hand
- symmetrical arthritis
- rheumatoid nodules
- serum RF
- radiological changes to the bone

Laboratory-based tests would include investigating the composition, particularly protein analyses, of blood and cerebrospinal fluid. Rheumatoid factor can be detected in plasma or serum using the Rose–Waaler test. In this test, sheep erythrocytes are coated with anti-sheep IgG at too low a concentration to cause agglutination. However, the addition of RF (remember it is an antibody) will result in agglutination (*Figure 5.21*).



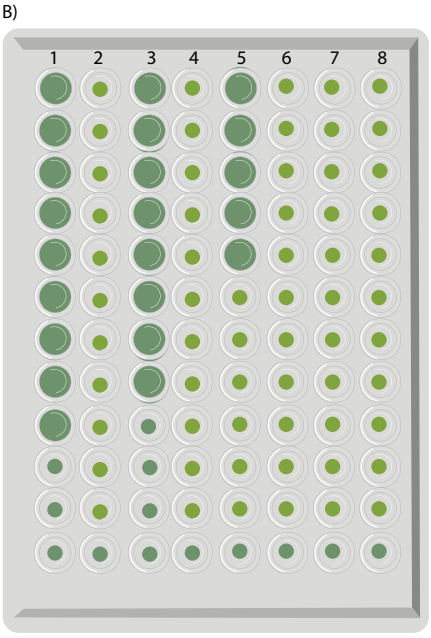


Figure 5.21 (A) Schematic illustrating the Rose–Waaler test for rheumatoid factor (RF), which measures the ability of a patient's serum to agglutinate sensitised sheep erythrocytes (commonly referred to as sheep red blood cells [SRBCs]). (B) The result of the test on several sera. Patients 1, 3 and 5 show positive agglutination at high dilutions, indicating the presence of RF.

Alternatively, latex particles nonspecifically coated with IgG can be used. A confirmed diagnosis of RA may require observations and testing lasting four to six weeks. The anti-cyclic citrullinated peptide (anti-CCP) test measures anti-CCP antibodies produced by the immune system. Patients who show a positive test have an increased likelihood of developing RA, although not all RA patients express these antibodies. However, positive tests for both RF and anti-CCP antibodies mean a greater likelihood of a severe condition requiring intensive treatment.

Treatment of RA is largely aimed at reducing the inflammation and associated pain and mitigating long-term damage. It includes using immunosuppressive agents such as methotrexate (*Chapter 6*) and anti-inflammatory drugs, including steroidal anti-inflammatory drugs (SAIDs), such as corticosteroids, and the nonsteroidal anti-inflammatory drugs (NSAIDs), for example large doses of aspirin. The long-term use of immunosuppressive drugs leads to an increased susceptibility to infections. More potent treatments are the disease-modifying antirheumatic drugs (DMARDs), for example hydroxychloroquine. If started sufficiently early, these therapeutic agents can be effective in managing the condition. The prognosis for patients depends on how early treatment is started, otherwise this debilitating condition can result in severe joint damage, deformity and loss of mobility.

Systemic lupus erythematosus

Systemic lupus erythematosus (SLE) is a chronic autoimmune disease of uncertain aetiology. The condition primarily affects women of reproductive age, although it can affect both much younger and much older patients. The female to male ratio is about 4:1, although in younger patients this gender bias does not occur. Initial presentations of SLE are arthritis and skin rash, particularly a 'butterfly' rash on the face that can be photosensitive (*Figure 5.22)*; however, the clinical features are extensive:

- symmetrical arthritis involving any joint
- arthralgia (bone pain)
- erythematous rash; butterfly rash on face (*Figure 5.22*)
- mucosal ulcerations
- pleurisy
- pericarditis
- renal defects
- fever
- nervous system involvement (psychosis, depression, convulsions, migraine)
- heart disease
- eye defects (retinal vasculitis; corneal ulceration)
- gastrointestinal ulcers
- pancreatitis
- hepatitis
- Sjögren disease (involving autoimmune destruction of lacrimal and salivary glands)
- haemolytic anaemia

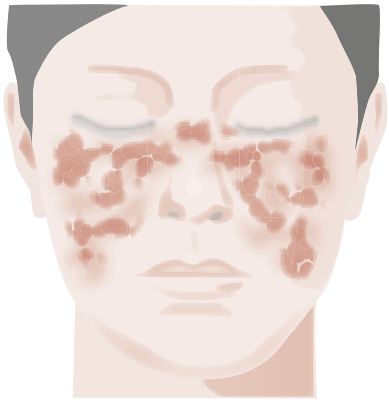


Figure 5.22 Schematic illustrating the area of face typically covered with a 'butterfly' rash in systemic lupus erythematosus (SLE).

Renal disease occurs in approximately half of affected individuals, with an early onset of nephritis.

About 95% of patients have antinuclear antibodies (ANAs) in their plasma. Other antibodies to erythrocytes, platelets and clotting factors may also be apparent, leading to the widespread deposition of immune complexes in tissues and inflammation that affects many body systems.

Unfortunately, the different patterns of symptoms displayed by patients have, in the past, presented difficulties with diagnosis. This systemic presentation is linked to the types of autoantibodies present. Diagnosis can be helped by demonstrating the presence of ANAs in samples of plasma, using indirect immunofluorescence (IDIF) on cultured Hep-2 cells. The most common pattern seen is a diffuse staining throughout the nucleus due to antibodies against chromatin. It is also possible to detect different patterns of fluorescence that are indicative of antibodies to different nuclear antigens, a feature that can help in diagnosis or in assessing prognosis (*Table 5.6*). Autoantibodies against **extractable nuclear antigens** (ENAs) can also be detected by radioimmunoassay (RIA) or ELISA (*Chapter 4*). Positive tests for ANAs require further investigations, for example for the presence in the blood of so-called lupus erythematosus (LE). These are neutrophils containing phagocytised nuclei and resemble large multinucleate cells that are also indicative of SLE. Some of the other laboratory tests used in diagnosing SLE are listed in *Table 5.7*. Rheumatoid factor and the presence of antibodies against cardiolipin may also be present, although these are not specific for SLE.

Pattern of staining by indirect immunofluorescence (IDIF)	Target antigen	Use
Homogeneous ANA	Chromatin	Screening but present in some nonpatients
	Double-stranded DNA	Specific for SLE
Speckled ANA (coarse)	Sm[a]	Specific for SLE
	U1-RNP	Found in SLE and mixed connective tissue disease

[a] Antibodies to the Sm (Smith) antigen were first discovered in the serum of a patient with SLE called Stephanie Smith. These antibodies bind to molecules in the nucleus called small nuclear riboproteins.

Table 5.6 Antinuclear antibodies (ANAs) in systemic lupus erythematosus (SLE). ⏎

Laboratory test	Reasons
Full blood counts	To detect anaemia, low platelet counts (*Chapter 13*)
Creatinine and electrolytes	To detect damage to kidneys (*Chapter 8*)
Erythrocyte sedimentation rate (*Chapter 13*)	Rate increases in SLE
Urine tests	To detect proteinuria (*Chapter 8*)
Concentrations of complement proteins (*Chapter 4*)	C2 or C4 deficiencies predispose to SLE; inflammatory process in active SLE reduces complement concentrations

Table 5.7 Examples of laboratory tests for systemic lupus erythematosus (SLE). ⏎

Mild cases of SLE may respond to NSAIDs. Antimalarials are used to control skin problems and arthralgia (joint pain). Immunosuppressive drugs (*Chapter 6*) are often used, although care must be deployed in patients prone to kidney disease. In addition, patients maintained on immunotherapy are more susceptible to infectious diseases. The prognosis for sufferers of SLE depends upon the location of inflammation and its severity. It improves with early detection and treatment but remains poor for those patients with severe complications of the heart and kidney and neurological complaints. Nevertheless, the prognosis for patients has improved greatly since the 1950s, when most patients died within 10 years of diagnosis.

5.5 IMMUNOLOGICAL HYPERSENSITIVITIES

Immunological hypersensitivities are disorders in which immune responses to foreign immunogens result in tissue damage. They are the most common group of immunological disorders and, collectively, affect approximately 10% of the population. Although **hypersensitivity** implies an overreaction, there is essentially nothing abnormal about these immune responses. Rather, it is the extent and nature of the exposure to the immunogen that results in the clinical damage. Indeed, immunogens that provoke immunological hypersensitivities are often 'harmless' in that they are not necessarily infectious agents.

The term **allergy**, which originally meant 'altered reactivity', is sometimes used synonymously with hypersensitivity. Allergies are, however, only some of the specific types of hypersensitivity. In 1963, Philip Gell and Robin Coombs classified hypersensitivities into four 'types' based on the part of the immune system that caused the tissue damage (*Table 5.8*). This classification scheme is still used today, although not all hypersensitivities belong exclusively to one type. For example, immunological reactions to drugs can be both types I and III, while intrinsic allergic alveolitis has components of types III and IV. In addition, expansion of the classification to include further types has been suggested. For simplicity, this chapter retains their original classification.

Type	Comment	Examples of disorder	Immune system component involved
I	Immediate, anaphylactic	Allergic asthma, hay fever, food allergies, anaphylactic shock	IgE
II	Cytotoxic	Transfusion reactions, haemolytic disease of the newborn (HDN)	IgM, IgG and complement
III	Complex-mediated	Intrinsic allergic alveolitis, serum sickness	Antigen–antibody complexes (usually involving IgG)
IV	Delayed-type hypersensitivity (DTH)	Mantoux reaction, contact hypersensitivity	Sensitised CD4+ T lymphocytes

Ig, immunoglobulin.

Table 5.8 The Gell and Coombs classification of immunological hypersensitivities.

Type I hypersensitivities

Type I are the most common immunological hypersensitivities and include well-known conditions, such as allergic asthma, allergic rhinitis and food allergies (*Table 5.9*). They are relatively common conditions, estimated to affect 1 in 5–10 people. Type I hypersensitivities are also referred to as *immediate* because their effects become apparent within eight hours of exposure to the immunogen. The effects can be *relatively* minor, as in hay fever, or life threatening, as in atopic (*Margin Note 5.7*) or allergic asthma or anaphylactic shock (see *Box 5.7*).

Margin Note 5.7
Atopy

The term **atopy** refers to an inherited tendency to develop allergic conditions exemplified by asthma, rhinitis and eczema, that is, type I hypersensitivities. People who suffer type I hypersensitivities are frequently referred to as being atopic. There is a definite genetic predisposition to atopy, with different members of a family sometimes showing different manifestations.

Disorder	Effects
Allergic asthma	Severe inflammation in respiratory tract, severe respiratory distress
Allergic rhinitis (hay fever)	Seen in upper respiratory tract; excess mucus, sneezing and wheezing
Food allergies	Skin rash (hives), gastrointestinal 'cramps' and diarrhoea; may sometimes result in anaphylactic shock
Anaphylactic shock	Sudden drop in blood pressure, respiratory distress, skin rash, gastrointestinal 'cramps' and diarrhoea; may result in death within an hour of exposure
Drug allergies, for example to penicillin, sulphonamides, salicylates	May be trivial (as in a skin rash) or severe, as in anaphylactic shock (see *Box 5.7*)

Table 5.9 Examples and some of the effects of type I hypersensitivities.

The term **allergy** is often used for type I, although this term is also used for some other types of hypersensitivities. Immunogens that cause allergies are often referred to as **allergens**, and this is the term used here. Thus, the different manifestations (*Table 5.9*) of type I hypersensitivities depend on the degree of previous exposures to the allergen and also on the route of exposure.

The underlying cause is the production of IgE in response to the allergen. This type of antibody stimulates inflammatory responses that are aimed at eliminating parasitic worms (*Chapter 4*). Atopic individuals produce IgE in response to allergens that, in non-allergic individuals, would stimulate the production of IgG. Mast cells in tissues (and basophils of blood) have receptors for the Fc region of IgE, so that they bind IgE to their surfaces. The more sensitised an individual, the more IgE coats these cells. Further exposure to the same allergen results in its binding to the mast cell (*Figure 5.23*). This triggers an explosive degranulation of the mast cell with the release of pharmacologically active mediators. These include histamine (see *Figure 4.14*), which stimulates vasodilation,

smooth muscle contraction and mucus secretion, depending on the site in the body. The subsequent release of further mediators, for example leukotrienes and prostaglandins, which are synthesised at the cell membrane, promotes inflammation and smooth muscle contraction. These responses, which evolved as a defence against multicellular parasites (*Chapters 2* and *4*), cause the characteristic symptoms of the hypersensitivity.

The most appropriate strategy for treating type I hypersensitivities is to identify the allergens so that the sufferer can avoid them. Allergens can be identified by skin testing (see *Box 11.6*). Extracts of common allergens are injected intradermally and, in sensitised individuals, causative allergens will produce a 'wheal and flare' skin reaction within 20 minutes. The wheal is a raised red lump, while the flare is the surrounding red inflamed area. Type I hypersensitivity investigations include the radioallergosorbent test (RAST) that measures the level of allergen-specific IgE in the blood samples. This involves incubating samples of serum from a patient with the potential allergen immobilised on a solid support; any IgE that binds to the allergen is then detected by the addition of ^{125}I-labelled anti-IgE antibodies.

When complete avoidance of the allergen is not feasible, drugs may be used to control the symptoms. These include antihistamines, such as brompheniramine maleate and loratadine, anti-inflammatory drugs, such as corticosteroids, and sodium cromoglycate, which prevents mast cell degranulation.

Allergic asthma

Allergic asthma is a chronic inflammatory obstructive respiratory disorder. In 2023, the National Institute for Health and Care Excellence (NICE, UK) reported its incidence as 300 million sufferers globally, while Asthma UK has reported the annual number of deaths worldwide due to asthma to be 380 000–420 000. It is estimated that the UK has over eight million patients, with 160 000 people diagnosed yearly, leading to approximately 60 000 hospital admissions and 1000 related deaths. The incidence of asthma is increasing but, although the condition is atopic, with an inherited tendency, its aetiology is poorly understood.

Allergic asthma can cause a long-term decline in lung function that is life threatening. Patients are sensitised to a range of airborne allergens that activate mast cells, as described above, and cause increased immune reactions in the bronchial tree. This stimulates inflammation in the respiratory tract, narrowing the airways and eventually leading to hyperreactivity of bronchial tree muscles. Allergic asthma produces recurring episodes of paroxysmal **dyspnoea** (shortness of breath), which is often described as an intense tightening of the chest and a feeling of suffocation, difficulties in breathing caused by a narrowing of the airways that produces a characteristic wheezing or high-pitched whistling sound in the chest when inhaling and/or exhaling, coughing, and the bronchial secretion of viscous mucus. Asthmatic attacks are exacerbated by, for example, chest infections, obesity (*Chapter 10*), smoking, panic attacks, chronic obstructive pulmonary disease (COPD; see *Chapter 9*) and lung cancer (*Chapter 17*). The severity of the condition varies between patients, as does the frequency of the attacks. Allergic asthma is an extremely distressing condition, particularly during an attack, when severe respiratory distress may require emergency treatment.

The standard diagnostic tools to diagnose allergic asthma are an assessment of the above symptoms and pulmonary functions tests by **spirometry** (see *Box 5.6*). However, spirometry has several limitations in that results may be normal between asthmatic attacks and vital capacity is normally unaffected or only marginally reduced. Peak expiratory flow (PEF) should be measured daily (at home) to assess whether lung function is decreasing. Values decrease during asthma attacks. A value of only 80% of normal should alert the patient to initiate home medication plans, such as increasing the dose or frequency of β agonists and inhaling corticosteroids (see below). A value of 50% of normal should be considered a medical emergency and hospital treatment immediately sought. Chest X-rays may be used to determine mucus plugging in the lungs, and blood tests, particularly to assess eosinophil counts and concentrations of IgE, can assist diagnosis.

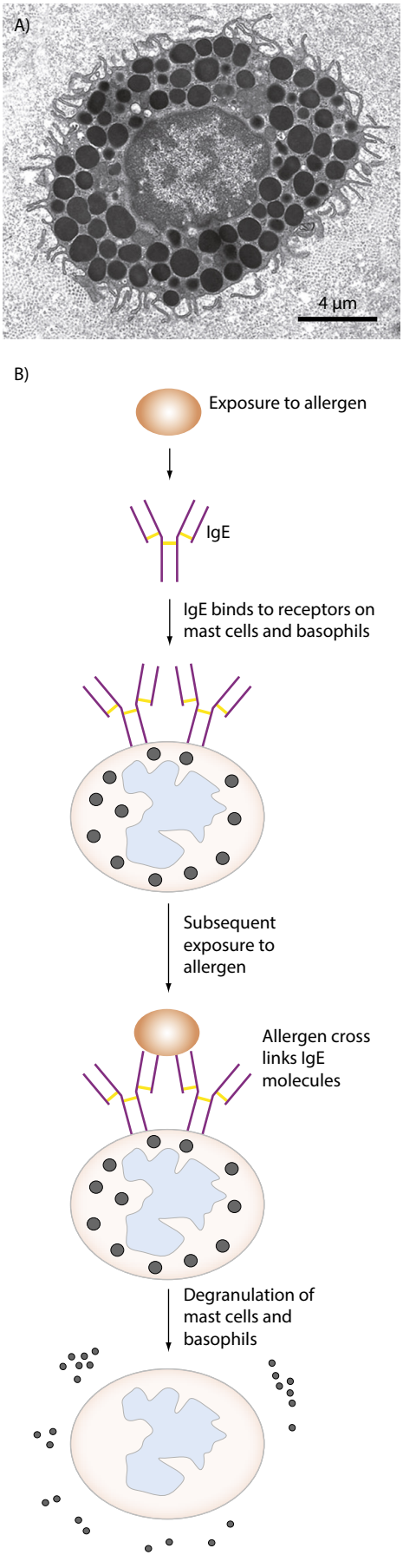


Figure 5.23 (A) Electron micrograph of a mast cell. Courtesy of Professor D.W. Vaughan, Boston University School of Medicine, USA. (B) Schematic illustrating how exposure to an allergen can lead to degranulation of mast cells and basophils and a type I hypersensitivity. IgE, immunoglobulin E. ⏎

BOX 5.6 LUNG VOLUMES AND PULMONARY FUNCTION TESTS

Determining lung volumes is an integral part of pulmonary function tests. These volumes (*Figure 5.24*) vary according to gender, age, ethnicity, body composition and respiratory diseases, such as asthma and chronic obstructive pulmonary disease (COPD; *Chapter 9*).

The combined volume of the lungs or *total lung capacity* (TLC) in adult human males and females is approximately 6 and 4 L, respectively. *Vital capacity* (VC) is the maximum volume of air that can be exchanged. In an adult male, this is about 4500 mL, leaving a *residual volume* (RV) of approximately 1500 mL that is never directly ventilated. The *functional residual capacity* (FRC) is the volume of air remaining in the lungs following a normal exhalation. Its normal value in adult males is about 3000 mL. At rest, a *tidal volume* (TV) of 300–500 mL, approximately 10% of VC, is inhaled or exhaled by normal adults during a respiratory cycle. However, no more than about 350 mL of the TV ever reaches the alveoli. The *inspiratory reserve volume* (IRV) is the volume that can be forcibly inhaled during deep breathing and forms a reserve to TV breathing; its normal adult male value is 1600 mL. The volume of air that can be forcibly exhaled by normal adult males, the *expiratory reserve volume* (ERV), is approximately 2500 mL. The inspiratory capacity (IC) is the volume, about 3000 mL, breathed in from quiet expiration to maximum inspiration. The relationships between these volumes can summed up as:

$$\begin{aligned} TLC &= RV + TV + IRV + ERV \\ VC &= TV + IRV + ERV \\ IC &= TV + IRV \\ FRC &= RV + ERV \end{aligned}$$

Tidal volume, IRV and ERV values can be measured by spirometry. However, determining RV, FRC and TLC requires other techniques. For example, TLC can be determined by helium dilution. Following quiet expiration, the subject breathes in a gas containing a known concentration of the inert gas helium. The subject then holds his/her breath for 10 seconds, which allows the helium to evenly diffuse into the air in the lungs. This dilutes the concentration of helium. The volume of air expired is measured and its concentration of helium determined. Total lung capacity can then be calculated. However, this value is accurate only if the lungs are not obstructed, since any obstruction will prevent the helium reaching all parts of the lung, producing an underestimate, as only *ventilated* lung volumes are measured.

Peak expiratory flow rate is a simple measure of the maximum flow of air that can be expelled from the lungs following full inspiration. The subject can measure it at home using a handheld meter as shown in *Figure 5.25*. The meter is reusable and simple to use: the subject simply takes as deep a breath as possible and then expires as hard and fast as possible into the meter.

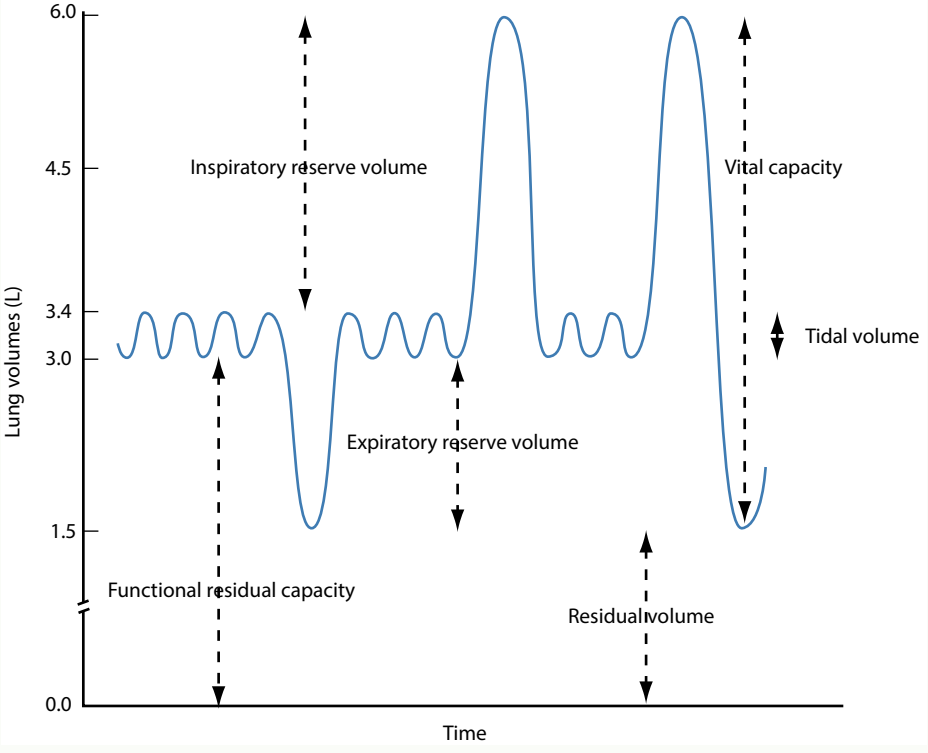


Figure 5.24 A stylised spirometry scan illustrating the lung volumes for a normal adult male.

Figure 5.25 (A) A peak flow meter. (B) The use of a peak flow meter. Courtesy of Asthma UK.

Figure 5.26 An inhaler of the type used in asthma prophylaxis and treatment. Courtesy of Asthma UK.

Treatment of allergic asthma includes avoidance of known allergen(s) and stimuli that precipitate attacks, such as cigarette smoke. Daily symptoms may be managed by inhalation (*Figure 5.26*) of salbutamol, a β agonist, sometimes with an added corticosteroid that relieves constriction of the airways. Inhalants are effective anti-inflammatory treatments: inhalation is more immediately effective than oral treatment, since the medication is delivered directly to the required body site. However, patients must always keep their inhalers with them. Treatment for severe acute attacks that need hospitalisation involves nebulised salbutamol with high doses of oxygen and injections of corticosteroids. Severe attacks refractory to drug therapy can lead to *status asthmaticus* and result in respiratory failure. This condition requires endotracheal intubation to save the patient.

Allergic rhinitis or hay fever

The effects of allergic rhinitis or hay fever are well known: a 'runny/blocked' nose; sneezing and coughing; itchy red eyes; congested throat, nose and ears; headaches; and a feeling of tiredness. The symptoms of asthma patients (see above) are also likely to be exacerbated. Allergic rhinitis is not life threatening but it can be very debilitating.

Seasonal allergic rhinitis typically presents in the spring and summer and is generally brought on by pollen (*Figure 5.27*) to which the individual has been sensitised. Some sufferers are sensitive to many types of pollen, including well-known 'culprits' such as ragweed, while others are allergic to a single type of tree pollen. Individuals who suffer allergic rhinitis all year round are most likely to be allergic to other types of allergens, such as the dead skin cells sloughed from household pets or the faeces of house dust mites (*Dermatophagoides pteronyssinus*), which thrive in warm, carpeted dwellings (*Figure 5.28*). There is no cure for allergic rhinitis but effective antihistamine medications to relieve the symptoms are available.

Figure 5.27 Photograph showing the release of pollen grains from Timothy grass (*Phleum pratense*). Courtesy of M. Hoult, Nature Photography, www.naturalhistoryimages.co.uk.

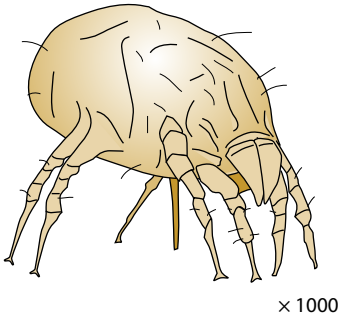


Figure 5.28 The house dust mite, *Dermatophagoides pteronyssinus*.

Food allergies

Allergies to some food are also seen in atopic individuals. Normally, the allergy will present as a skin rash, possibly with diarrhoea, within an hour of consumption of the food involved. Foods known to cause allergies include eggs, shellfish, mushrooms and strawberries. Recently, much attention has focused on the severe allergic reactions to nuts, especially peanuts. This may be due to the use of peanut paste as a thickening agent in the preparation of certain processed foods. Individuals who are allergic to peanuts may suffer rapid and life-threatening allergic reactions and are advised to avoid processed food unless the ingredients are clearly labelled. It has also been known for individuals to suffer severe reactions to certain fruits, such as papaya, although this is a rare condition.

BOX 5.7 ANAPHYLACTIC SHOCK ⏎

Anaphylactic shock is a life-threatening reaction to an allergen that may begin within minutes of exposure. It is often due to intramuscular or intravenous exposure to an immunogen, such as in a penicillin-sensitive individual who has been given a penicillin injection to treat a bacterial infection. People who are highly sensitised to bee venom (*Figure 5.29*) may also develop anaphylactic shock following a bee sting. Anaphylactic shock may also occur following sensitisation to the latex in some clinical gloves. In addition, food allergies in highly sensitised individuals occasionally result in anaphylactic shock.

Typically, anaphylactic shocks involve extreme respiratory distress, accompanied by hypotension, **urticaria** (skin rashes) and diarrhoea. The condition arises from a systemic release of mast cell mediators. The extensive vasodilation causes a sudden massive drop in blood pressure. The effects of mediators, in particular leukotrienes, on respiratory muscles also leads to breathing difficulties. Anaphylactic shock is a medical emergency and must be treated immediately with, for example, rapid intramuscular injection of adrenaline to relax the smooth muscle tissues.

Figure 5.29 The honeybee *Apis mellifera*. Courtesy of M. Hoult, Nature Photography, www.naturalhistoryimages.co.uk. ⏎

Type II hypersensitivities

Type II or cytotoxic hypersensitivity refers to situations in which an antibody activates complement, causing tissue damage. Examples include transfusion reactions involving mismatched bloods and haemolytic disease of the newborn (HDN; see *Chapter 6*). Autoimmune diseases (*Section 5.4*) that involve lysis of cells can also be regarded as type II hypersensitivities. Blood transfusion and HDN are described in *Chapter 6*; hence, only a brief discussion of the causes and consequences of the transfusion of ABO-incompatible bloods will be given here.

Transfusion reactions

The ABO blood group system contains four blood groups: A, B, AB and O. The antigens for these groups are different glycans (oligosaccharides), as shown in *Figure 5.30*, carried on erythrocyte membranes (*Chapter 6*). Plasma contains antibodies to the blood group antigens: individuals with blood group A have anti-B antibodies and those of group B have anti-A. These antibodies, known as **isohaemagglutinins**, usually belong to the IgM class, which are efficient activators of complement. If a blood group A individual is and transfused with group B blood, then antibodies from the donated blood will bind to recipient erythrocytes, activate complement and cause their lysis. Similarly, antibodies in the recipient will lyse the donated erythrocytes (*Figure 5.31*). The sudden and simultaneous lysis of cells leads to kidney failure and possible death.

Type III hypersensitivities

Type III or complex-mediated hypersensitivities arise from immune complexes that usually involve antibodies to soluble antigens. Immune complexes can be harmful because they activate complement, triggering inflammation and the influx of neutrophils into an area. Over time, this can lead to tissue damage, principally by lytic enzymes released from dying neutrophils. The sizes of immune complexes vary in accordance with the relative proportions of antigen to antibody; hence, the clinical consequences also vary.

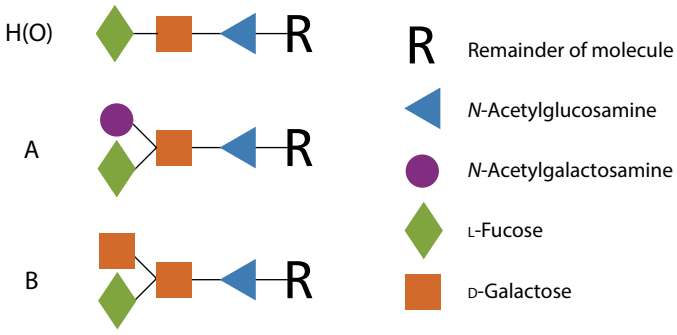


Figure 5.30 Cartoon illustrating the structures of the ABO blood group antigens.

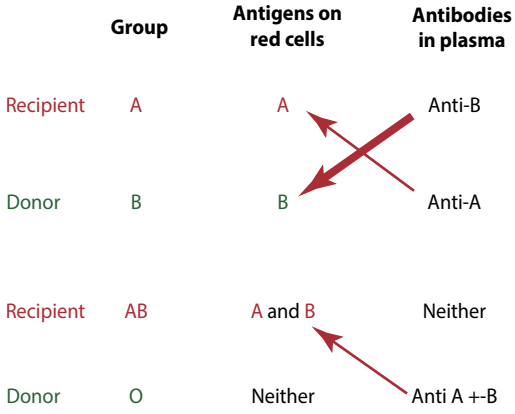

Group
Antigens on red cells
Antibodies in plasma
Recipient A A Anti-B
Donor B B Anti-A
Recipient AB A and B Neither
Donor O Neither Anti A +-B

Figure 5.31 The antigens and antibodies of the ABO blood group system. The results of only two incompatible transfusions are presented, with the thickness of the arrows indicating the relative amounts of agglutination (see *Chapter 6*).

For example, in RA (*Section 5.4*), immune complexes formed between the RF and IgG are produced in an excess of antigen and are small and soluble and travel in the circulation. They may adhere to the insides of blood vessels triggering vasculitis or terminate in the kidney and cause nephritis. However, in extrinsic allergic alveolitis, immune complexes produced in the alveoli are large and precipitate in the lungs, causing alveolitis.

The term 'intrinsic allergic alveolitis' covers a number of occupational diseases that result from exposure to airborne immunogens. A typical example is farmer's lung (*Figure 5.32*), which typically begins in winter when the farmer develops a cough. Since this is not an unusual occurrence, the disease may go unrecognised for several years. However, the disease is progressive and, if untreated, will progress to emphysema. The disease is related to the farmer moving hay to feed livestock in winter. Ascomycete fungi grow well in the warm damp conditions in stored hay, and clouds of fungal spores are released when it is moved. Inhalation of the spores over months and years results in large concentrations of circulating IgG to immunogenic molecules leached from the spores in the farmer. Further exposure to spore antigens results in large immune complexes of spore antigen–IgG precipitating in the lungs, causing inflammation within the alveoli. Successive winters may result in immune complex-mediated damage to the lungs, with fibrosis and loss of gas exchange capacity.

Although farmer's lung is classified as an antibody-mediated hypersensitivity, it is now recognised that cell-mediated immunity is also involved and that damage caused by specific T cells also contributes to the disease.

It is essential that farmer's lung be diagnosed early to avoid permanently damaged lung tissue. Treatment in the early stages may simply involve avoidance of the antigen, although corticosteroids may also be used to treat the inflammation.

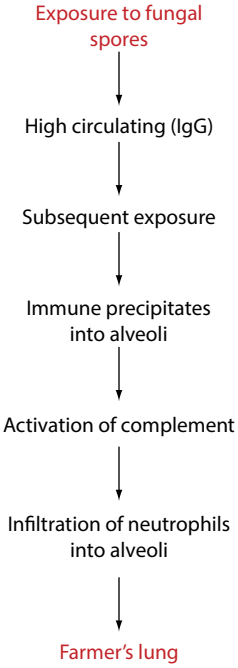


Figure 5.32 A long-term exposure to spores can lead to high concentrations of circulating antibodies (immunoglobulin G, IgG) that can precipitate spores in the alveoli and activate complement, leading to farmer's lung.

Type IV delayed-type hypersensitivities

Type IV or delayed-type hypersensitivity (DTH) requires more than 18 hours for the symptoms to become apparent after exposure to an immunogen. This type of hypersensitivity is caused by T lymphocytes rather than antibodies. The **Mantoux reaction** is a typical DTH, which occurs following vaccination against tuberculosis (TB). The vaccine consists of an attenuated form of Bacillus Calmette–Guérin (BCG), which is a strain of *Mycobacterium bovis* (*Chapter 2*). Individuals are skin tested before being given the vaccine to determine if they have already been sensitised by contact with *M. bovis* or by a previous vaccination. The skin test involves injecting an extract of mycobacteria, called **purified protein derivative** (PPD), intradermally. In sensitised individuals, the injection site becomes swollen and red after 18–24 hours. Swelling increases for approximately 48 hours but then subsides slowly, so that this 'positive' reaction is still visible after several weeks. An individual who gives a positive Mantoux test may then be investigated to ensure that the positive result was due to prior sensitisation rather than active disease! Initially, in a positive test, sensitised CD4+ T cells respond to PPD by releasing cytokines that attract small lymphocytes and monocytes, which form the swelling at the injection site (*Figure 5.33*). This induces inflammation, allowing the entry of more CD4+ cells.

Thus, a cascade reaction occurs, producing a slow but progressive swelling after the injection. The reaction subsides once the phagocytic monocytes have removed the PPD.

Delayed-type hypersensitivity causes **contact allergies** to a variety of materials, including certain biological stains, some hair dyes, nickel used in cheap jewellery, mercuric salts in some tattoo dyes, fluorodinitrobenzene and some plant biochemicals, such as urushiol in poison ivy. Typically, once sensitised, an individual will develop dermatitis 18–24 hours after *further* skin contact with the same material. Such skin-sensitising materials are not typical immunogens, since they are neither proteins nor large molecules. Sensitisation appears to involve their chemical binding to skin proteins to form hapten–protein conjugates that are immunogenic (*Chapter 4*). Langerhans cells, which are antigen-presenting cells in the skin, process this 'new' antigen and present the modified peptides to helper T lymphocytes, which become sensitised. Once this has happened, any further contact with the hapten–protein conjugates will promote cytokine release from T cells, producing a cascade similar to that of the Mantoux reaction.

Laboratory workers need to be aware of the skin-sensitising nature of many of the chemicals used in the biomedical sciences and conduct risk assessments for their uses, since sensitisation, once induced, is irreversible. The only treatment for type IV hypersensitivities is to avoid the allergen, although corticosteroid creams relieve the dermatitis.

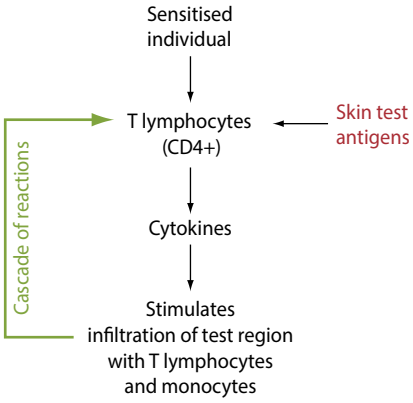


Figure 5.33 The Mantoux test for type IV hypersensitivity.

BOX 5.8 SOME CLINICAL FEATURES OF CYTOKINES

Cytokines and their soluble receptors are present in the body in low concentrations (*Chapter 4*), but many are now available in larger quantities thanks to molecular biological techniques, such as cloning. This, together with the use of monoclonal antibodies that can inhibit their actions, means a number of cytokines are now of clinical significance. Indeed, interferon (IFN)-α has been used to treat Kaposi sarcoma, hairy cell leukaemia (*Chapter 17*) and hepatitis B and C (*Chapters 2* and *3*), while IFN-β is used in the treatment of chronic granulomatous disease. Several immunodeficiency diseases are treated with recombinant granulocyte colony-stimulating factor (GCSF) and IFN-γ. Cytokine excesses are also involved in the pathology of a number of diseases, particularly inflammatory disorders, such as rheumatoid arthritis.

The uses of monoclonal antibodies, which prevent cytokine–receptor binding, and the application of soluble receptors that interfere with cytokine–membrane receptor interactions are also receiving clinical attention.

5.6 CASE STUDIES

CASE STUDY 5.1

George is a nine-month-old boy who has suffered repeated bouts of upper respiratory tract infections since the age of four months. He was admitted to hospital with bacterial pneumonia. A routine examination of his throat showed his tonsils to be much smaller than usual in boys of his age, especially those who have had numerous throat infections. George has no siblings.

Questions

(a) What diagnoses might plausibly be suggested?
(b) How could these suggestions be confirmed?
(c) If confirmed, what types of treatment would be beneficial?

CASE STUDY 5.2

Peter is a teacher at a primary school. In his spare time he keeps bees and sells the honey he produces at local country fairs. At one such fair he was allowing potential customers to sample his honey. The stall became very attractive to some nearby wasps and, in trying to keep the wasps away, Peter was stung. Within a few minutes he collapsed with obvious severe breathing problems and a generalised rash.

Questions

(a) What is the most likely cause of Peter's collapse?
(b) How should he be treated?

CASE STUDY 5.3

Jane is a 24-year-old student studying for a PhD. She is hard-working and dedicated and her first two years went well. In the last year, however, her attendance at university declined and her work suffered. Her supervisor was worried about her constant tiredness and weight gain and, although the university is well heated, Jane was always complaining about feeling cold. Her supervisor advised her to consult her physician, who notices that Jane has a goitre.

Questions

(a) What is the likely cause of Jane's tiredness?
(b) What clinical tests are appropriate for Jane?

CASE STUDY 5.4

Martha and John are a young couple who are desperate to start a family. Sadly, they had lost their first baby at the age of nine months following chronic lung infections. Thus, they were understandably concerned when their five-week-old baby, Gavin, developed a persistent temperature of 39°C, combined with inflammation at the umbilical cord stump. At the accident and emergency department, staff were very concerned for Gavin, especially given his clinical history of infections and the family's clinical history, and the family were retained in hospital for extensive clinical tests on Gavin. Blood tests showed that all classes of leukocytes were markedly above reference values. A Rebuck skin window test showed a lack of neutrophil and monocyte mobility even after eight hours. Flow cytometry indicated a deficiency in CD18+ T cells.

Questions

(a) What does the fever and increased numbers of leukocytes indicate?
(b) What can be concluded about the immobility of neutrophils and monocytes and absence of CD18+ T cells?
(c) What is the most plausible diagnosis of Gavin's condition?
(d) Given this diagnosis, how should he be treated?

CASE STUDY 5.5

Lily, a 15-year-old comprehensive school student, attended an antiques emporium with her mother where she bought a 1920s costume jewellery set of matching necklace and earrings in a style that was fashionably 'retro'. A week later she wore the jewellery for the first time at her elder sister's wedding. A fortnight after that, she wore them again at a sleepover birthday party. Unfortunately, the following evening her ear lobes and neck became very swollen, inflamed and painful.

Questions

(a) How would you describe Lily's condition?
(b) What has most likely caused this clinical problem?
(c) How should Lily's condition be treated?

5.7 SUMMARY

Disorders of the immune system include immunodeficiencies, autoimmune diseases and hypersensitivities. The lack of a component of the immune system renders the sufferer much more susceptible to infectious diseases and some forms of cancer. These diseases may be treated with antibiotics and antiviral drugs to treat infections and, where possible, to replace the missing component by a BMT/HSCT. Autoimmune diseases occur when mechanisms for preventing self-reactivity fail. These diseases can be organ specific or systemic. They are treated palliatively to relieve symptoms. Immunosuppressive drugs are used, although this renders the patient more susceptible to infections. Hypersensitivities are classified into four types depending on the underlying cause. They are not curable and best treated by avoiding contact with the immunogen to which the individual is sensitised; however, medications are available to alleviate their symptoms.

QUESTIONS

1. Which ONE of the following statements is CORRECT?
(a) Rheumatoid factor is only present in patients with RA.
(b) All patients with RA have RF.
(c) Antibodies to double-stranded DNA are diagnostic for SLE.
(d) Graves' disease is characterised by antibodies to the acetylcholine receptor.
(e) Goodpasture disease is a systemic disorder.

2. Which of the following is the odd ONE out?
(a) $T^-B^-NK^-$.
(b) B and T lymphocytes.
(c) *STAT3*.
(d) *ADA*.
(e) *RAG1*.

3. State whether the following statements are TRUE or FALSE.
(a) Bruton agammaglobulinaemia is an autosomal recessive condition.
(b) Complement deficiencies result in increased neisserial infections.
(c) Deficiencies in T cells cause an increased susceptibility to viral infections.
(d) T cell deficiencies are corrected with infusions of plasma.
(e) Immunodeficiency may lead to increased risk of some cancers.
(f) Antiglomerular basement membrane disease is an autoimmune disease and predominantly affects females.
(g) The immunoglobulin IgD is a pentamer.
(h) The sialyl Lewis X glycan is also known as CD15.
(i) Chronic granulomatous disease can be both X- and autosomally linked.
(j) Mutations in *ITGB2* at location 22q21.3, X-linked means macrophages, neutrophils and lymphocytes do not express β_2 integrin.

4. Which of the following statements is INCORRECT?
(a) In autoimmune disease generally, more women than men are affected.
(b) Transient hypogammaglobulinaemia refers to a brief antibody deficiency in pregnant women.
(c) Deficiencies of C9 are asymptomatic.
(d) The classical pathway for complement action depends on the presence of antibodies.
(e) In SCID, B lymphocyte levels may be normal or decreased.

5. List three ways in which the body prevents autoimmune reactions.

6. Arrange the following two lists into their most appropriate pairings

Selective	SLE
Farmer's lung	Acetylcholine receptors
Chédiak–Higashi disease	May be asymptomatic
Anaphylactic shock	IgA deficiency is due to IgE
Delayed hypersensitivity	Is caused by immune complexes
HSCT	Possible GVHD
XLA	BTK
ANAs	Affects melanocytes
Myasthenia gravis	Is caused by T lymphocytes
IgM immunoglobulin	RF

7. Using the following equations, complete *Table 5.10*:

TLC = RV + TV + IRV + ERV
VC = TV + IRV + ERV
IC = TV + IRV
FRC = RV + ERV

	Lung volumes/mL
TLC	4000
TV	
RV	
IRV	1600
ERV	1200
FRC	
VC	
IC	2000

Table 5.10 ⏎

FURTHER READING

Abbas AK *et al*. (2019) *Basic Immunology, Functions and Disorders of the Immune System*, 6th edition. Elsevier.

Almarza Novoa EA *et al*. (2018) Leukocyte adhesion deficiency-I: a comprehensive review of all published cases. *J Allergy Clin Immunol Pract* **6**:1418–1420. e10.

Angum F *et al*. (2020) The prevalence of autoimmune disorders in women: a narrative review. *Cureus* **12**:e8094. doi: 10.7759/cureus.8094.

Anjani G *et al*. (2020) Recent advances in chronic granulomatous disease. *Genes Dis* **7**:84–92.

Arnold DE and Heimall JR (2017) A review of chronic granulomatous disease. *Adv Ther* **34**:2543–2557.

Azizi G *et al*. (2017) Abnormality of regulatory T cells in common variable immunodeficiency. *Cell Immunol* **315**:11–17.

Candotti F (2018) Clinical manifestations and pathophysiological mechanisms of the Wiscott–Aldrich syndrome. *J Clin Immunol* **38**:13–27.

Cavazzana M *et al*. (2019) Gene therapy targeting haematopoietic stem cells for inherited diseases: progress and challenges. *Nat Rev Drug Discov* **18**:447–462.

Cevhertas L *et al*. (2020) Advances and recent developments in asthma in 2020. *Allergy* **75**:3124–3146.

Darian T *et al*. (2020) A review of primary immune deficiency disorders. *EMJ Allergy Immunol* **5**:70–77.

Demirdag YY and Gupta S (2021) Update on infections in primary antibody deficiencies. *Front Immunol* **12**:634181.

Grumach AS and Kirschfink M (2014) Are complement deficiencies really rare? Overview on prevalence, clinical importance and modern diagnostic approach. *Mol Immunol* **61**:110–117.

Kaul A *et al*. (2016) Systemic lupus erythematosus. *Nat Rev Dis Primers* **2**:16039. doi: 10.1038/nrdp.2016.39.

Kuo CY *et al*. (2018) Immune and genetic features of the chromosome 22q11.2 deletion (DiGeorge syndrome). *Curr Allergy Asthma Rep* **18**:75–81.

Ling M and Murali M (2019) Analysis of the complement system in the clinical immunology laboratory. *Clin Lab Med* **39**:579–590.

Mortuaire G *et al*. (2017) Specific immunotherapy in allergic rhinitis. *Eur Ann Otorhinolaryngol Head Neck Dis* **134**:253–258.

Mulders-Manders CM and Simon A (2015) Hyper-IgD syndrome/mevalonate kinase deficiency: what is new? *Semin Immunopathol* **37**:371–376.

Műzes G and Sipos F (2018) Primary immunodeficiency and autoimmune diseases. *Orv Hetil* **159**:908–918.

Quirt J *et al*. (2018) Asthma. *Allergy Asthma Clin Immunol* **14**: (Supplement 2):50 doi: 10.1186/s13223-018-0279-0.

Rose NR (2017) Autoimmune diseases. In: *International Encyclopedia of Public Health*, 2nd Vol. eds Rose NR and Mackay IR. Springer.192–195.

Saito M *et al*. (2016) Molecular mechanisms of nickel allergy. *Int J Mol Sci* **17**:202. doi: 10.3390/ijms17020202.

Schröder-Braunstein J and Kirschfink M (2019) Complement deficiencies and dysregulation: pathophysiological consequences, modern analysis, and clinical management. *Mol Immunol* **114**:299–311.

Sharma P *et al*. (2020) Chediak–Higashi syndrome: a review of the past, present, and future. *Drug Discov Today Dis Models* **31**:31–36.

Tangye SG *et al*. (2020) Human inborn errors of immunity: 2019 update on the classification from the International Union of Immunological Societies Expert Committee. *J Clin Immunol* **40**:24–64.

Tao J-H *et al*. (2017) Foxp3, regulatory T cell, and autoimmune diseases. *Inflammation* **40**:328–339.

Thrasher AJ and Williams DA (2017) Evolving gene therapy in primary immunodeficiency. *Mol Ther* **25**:1132–1141.

van der Meer JWM and Simon A (2016) The challenge of autoinflammatory syndrome: with an emphasis on hyper-IgD syndrome. *Rheumatology (Oxford)* **55**:(Supplement 2): ii23–ii29.

Weber ANR *et al*. (2017) Bruton's tyrosine kinase: an emerging key player in innate immunity. *Front Immunol* **8**:1454.

Zimmermann SC *et al*. (2019) Dismantling airway disease with the use of new pulmonary function indices. *Eur Respir Rev* **28**(151):180122. doi: 10.1183/16000617.0122-2018.

Useful websites

http://www.ncbi.nlm.nih.gov/structure
https://www.omim.org
https://www.asthmaandlung.org.uk/
www.transfusionguidelines.org
https://emedicine.medscape.com/

Taylor & Francis
Taylor & Francis Group
http://taylorandfrancis.com

Transfusion and transplantation

6

OBJECTIVES

After studying this chapter you should be able to:

- outline the uses of transfused blood and blood products
- describe the nomenclature of blood groups
- describe the characteristics of the ABO, Rhesus and selected other blood groups
- outline ways in which blood groups are determined
- discuss some of the possible adverse effects of transfusion
- outline how tissue typing is performed
- outline the roles of the immune system in rejecting tissue/organ transplants
- discuss the significance of major histocompatibility complex (MHC)/human leukocyte antigen (HLA) proteins in transplant rejection
- outline the actions and uses of some immunosuppressive drugs
- outline graft versus host reactions
- explain the types and uses of haematopoietic stem cell transplants

6.1 INTRODUCTION

Blood transfusion is the transfer of blood or blood products into the cardiovascular system of a patient who has lost blood due to injury, disease or an operation (*Figure 6.1*). Transplantation is the transfer of biological material, cells, tissues and organs, from one position in the body to another or from one person (the donor) to a patient (the recipient). In that sense, blood transfusions are transplantations, but the term **transplant** usually refers to the donation of solid tissues/organs or haematopoietic stem cells (HSCs) harvested from blood or bone marrow.

During a transfusion, the volume and type of blood, or blood component, administered depends on the needs of the patient. Transfusions of blood and blood products are routine and are generally safe therapeutic procedures that are rarely associated with adverse reactions. However, it is only in the last 100-or-so years that the foundations of these procedures have been established. This chapter will explain the uses of blood transfusion and blood products, discuss the biochemical and genetic bases of several blood group systems, and describe the methods used to determine blood groups. In addition, some adverse reactions of transfusion and the consequences for a fetus if antibodies against fetal erythrocytes moving across the placenta will be outlined.

The transplantation of solid tissues and organs from one individual to another to replace diseased or nonfunctional tissues/organs is now well established. However, it is only about 70 years since the first successful kidney transplant in 1954 paved the way for the range of tissue/organ transplants currently available.

The genetics of the human leukocyte antigen (HLA) system, a series of genes encoding proteins that stimulate rejection of tissues/organ transplants, will be examined, together with the immunological bases of transplant rejection. This chapter will also consider how the rejection of transplanted tissues and organs can be prevented by immunosuppressive therapy, review the circumstances that result in **graft versus host disease** (GVHD) and consider the consequences of GVHD for treatments involving the transplantation of HSCs.

Figure 6.1 A blood bank and transfusion centre. Courtesy of the Manchester Blood Transfusion Service, UK.

DOI: 10.1201/9781003296836-6

6.2 BLOOD AND BLOOD PRODUCTS FOR TRANSFUSION

The major role of a transfusion laboratory (*Figure 6.1*) is to ensure that blood and blood products can safely be used to treat a patient. To ensure safety, the blood is tested to determine its blood group and to investigate whether it is contaminated with harmful viruses and/or microorganisms. In addition, checks are made to ensure that the transfused blood does not contain antibodies that will destroy the erythrocytes of the recipient and possibly cause death.

Blood transfusions are required to replace blood lost as a result of accident or surgery. Surgical procedures that require transfusions include the transplantation of organs, such as kidneys, livers, lungs and hearts, where significant bleeding may occur. Blood may also be given to treat certain diseases, such as anaemia (*Chapter 13*). Plasma may also be transfused to treat badly burned patients who have lost significant amounts of fluid or to treat bleeding disorders, as are platelet concentrates (*Figure 6.2*). Plasma products, such as factor VIII, necessary in the treatment of haemophiliacs (*Figure 6.3* and *Chapter 13*), or immunoglobulins to treat certain immunodeficiency disorders, may also be given (*Chapter 5*). In the UK, blood containing leukocytes is not transfused for several reasons (*Table 6.1*).

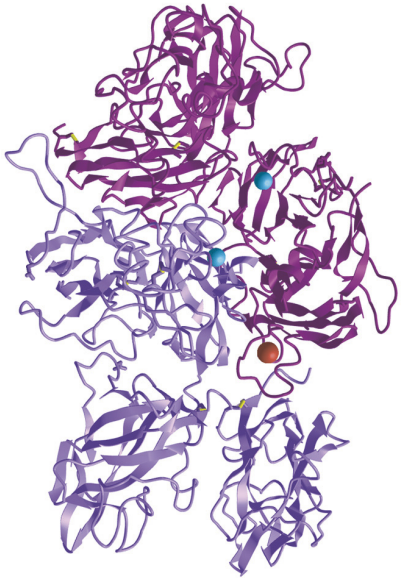


Figure 6.3 Molecular model of factor VIII. Its two Cu^{2+} and single Ca^{2+} are coloured blue and reddish-brown, respectively. PDB file 3CDZ. ⏎

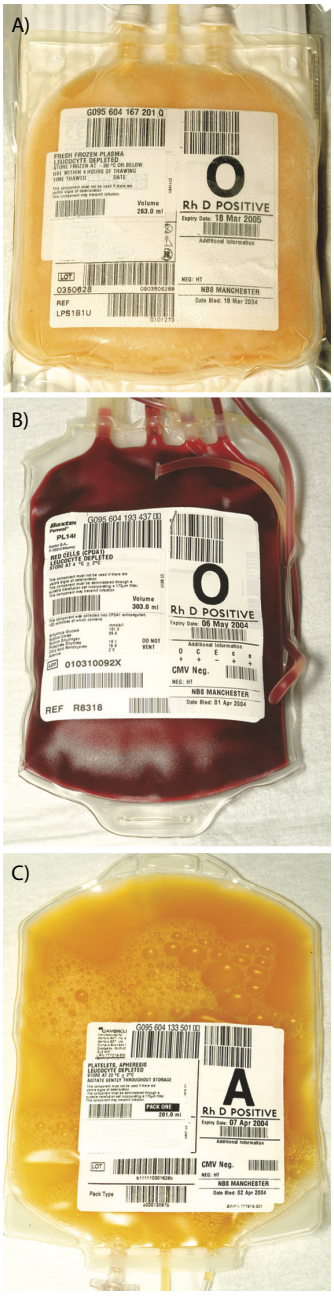


Figure 6.2 Some blood products: (A) plasma, (B) erythrocytes and (C) platelets. Courtesy of the Manchester Blood Transfusion Service, UK. ⏎

Reason	Examples
Preventing the possibility of transmitting infectious agents in leukocytes	Human immunodeficiency virus (HIV) Cytomegalovirus (CMV) Prion proteins responsible for variant Creutzfeldt–Jakob disease (vCJD)[a]
Reduces the possibility of immunising against leukocyte antigens	Sensitisation to major histocompatibility antigens
Decreases risk of graft versus host disease (GVHD) associated with whole blood	In immunosuppressed or immunodeficient individuals
Reduces the possibility of inducing febrile reactions post-transfusion (nonhaemolytic febrile reactions)	Caused by cytokines (*Chapter 4*) released from leukocytes during storage

[a] In 2003, a recipient of blood taken from a healthy donor who developed vCJD also died of vCJD six years after transfusion. Other evidence since that time suggests that it is possible to transmit vCJD with leukocytes (see also *Chapters 2* and *15*).

Table 6.1 Some of the reasons for leukodepleting blood before transfusion. ⏎

Sensitisation to major histocompatibility complex (MHC) or HLA (*Chapter 4*) antigens, which are present on leukocytes but not erythrocytes, may have consequences if the recipient later requires a transplant (*Section 6.11*) and GVHD may have a fatal outcome in immunosuppressed individuals. Accordingly, leukocytes are removed from blood, usually within a few hours of collection, by filtering the blood through leukocyte-specific filters. These trap leukocytes but not the smaller erythrocytes or platelets. The process, called **leukodepletion**, reduces the leukocyte count to less than 5×10^6/L (see *Box 5.2*). The number of leukocytes left in blood can be assessed by counting in a haemocytometer or by flow cytometry (*Box 6.1*).

BOX 6.1 FLOW CYTOMETRY

Figure 6.4 A flow cytometer.

A flow cytometer (*Figure 6.4*) can distinguish different cells from mixed populations by analysing several of their properties simultaneously. The cells to be analysed are passed individually in a stream past a laser light source. This is usually achieved by having a sheath of fluid passing through an orifice of 50–300 μm diameter. The sample is injected into the sheath fluid as it passes through the orifice. The sample and sheath fluid do not mix if appropriate fluid flow rates are chosen (*Figure 6.5*). The laser beam illuminates the cells, which scatter some of this light. The scattered light is detected simultaneously by two detectors: one measures forward scatter, that is, the light scattered up to 10° from the incident beam; the other detects side scatter, the light deflected 90° from the incident beam. The intensities of the forward and side scatters are related to the sizes and shapes of the cells. The side scatter arises more from the granularity of a cell, while the forward is sensitive to its surface characteristics. Thus, a mixed population of cells can be analysed on the bases of these measurements.

In addition to side and forward scatters, cells can be stained with fluorescent antibodies to proteins that are characteristic of a cell population. For example, in a mixed population of lymphocytes, the T lymphocytes (*Chapter 4*) could be stained with a fluorescent antibody to the CD3 marker. This allows the T to be distinguished from B lymphocytes that lack this protein. As the cells are illuminated in the laser beam, the fluorescent T lymphocytes can be detected by additional detectors using appropriate filters. The intensity of the fluorescence is related to the number of CD3 molecules on the cells (*Figure 6.6*). It is also possible to distinguish different

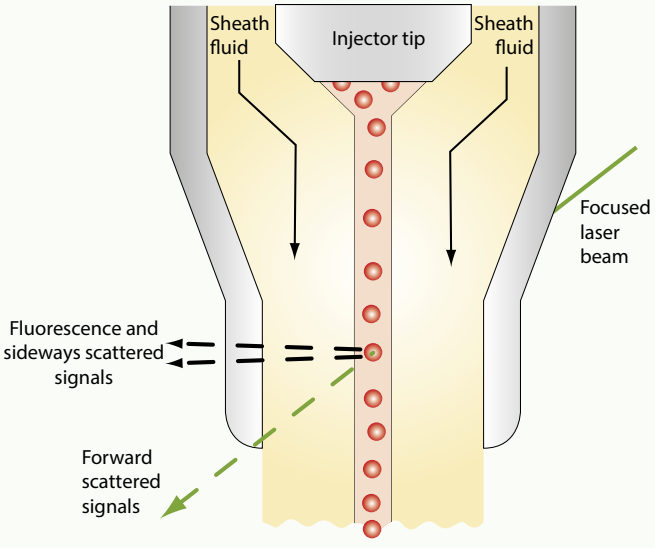


Figure 6.5 Schematic of a flow cell.

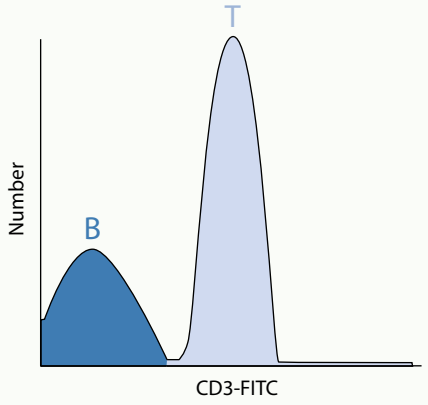


Figure 6.6 Diagram showing how T and B lymphocytes can be identified and distinguished by flow cytometry. FITC, fluorescein isothiocyanate.

populations simultaneously if they are stained with antibodies to unique marker proteins and if the antibodies are conjugated to different fluorescent dyes. Thus, T_H lymphocytes (*Chapter 4*) can be stained with a fluorochrome-labelled antibody to CD4. Simultaneously, *all* the T cells can be stained with an antibody, conjugated to a different fluorochrome, to CD3. Analysis by flow cytometry would reveal four populations of small lymphocytes (*Figure 6.7*): those which are CD3– and CD4– (B lymphocytes), those that are CD3+ and CD4– (presumably T_C lymphocytes), those that are CD3– and CD4+ (a very minor population) and those that are CD3+ and CD4+ (the T_H cells).

Flow cytometry is used in many aspects of pathology science. In transfusion it can be used, for example, to measure the number of fetal erythrocytes in the maternal circulation following a placental bleed. In transplantation it may be used to evaluate the results of a cross match (*Sections 6.11* and 6.12).

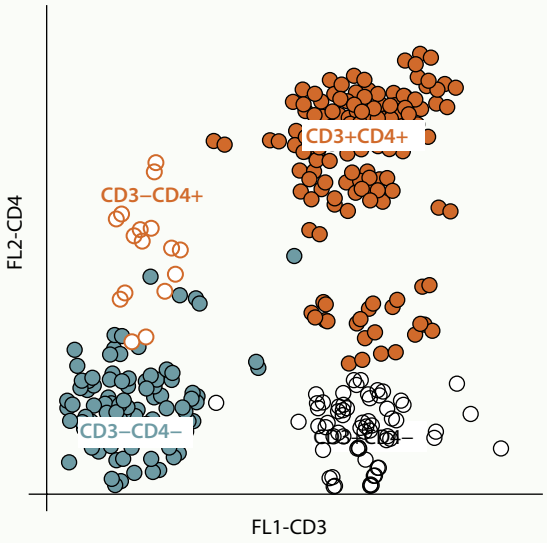


Figure 6.7 Schematic showing how the double staining of lymphocytes in flow cytometry can distinguish several different populations of cells. The cells are represented by different colours on the diagram.

6.3 THE DISCOVERY OF THE BLOOD GROUP SYSTEMS

It has been known since the seventeenth century that the transfusion of blood between individuals could have rapid and fatal consequences. Fortunately, in 1900, Karl Landsteiner (1868–1943) discovered that individuals could be classified into different groups depending on the characteristics of their erythrocytes. Landsteiner drew blood from a number of individuals and separated the erythrocytes from the plasma. He then mixed all possible combinations of erythrocytes and plasma from these individuals together and showed that only certain combinations resulted in the clumping or **agglutination** of erythrocytes (see *Figure 6.8*). These discoveries laid the foundations for the routine and safe therapeutic transfusion of blood. The patterns of agglutination showed that there were different blood groups, which Landsteiner named A, B and O. In 1902, Alfred von Decastello (1872–1960) and Adriano Sturli (1873–1964) discovered a fourth blood group, which was called AB. It became clear that fatal blood transfusions resulted from incompatible bloods being transfused and that these procedures were successful when the blood transfused was of an identical blood group. Landsteiner later discovered other blood group systems, including the Rhesus (Rh) system. Since then, numerous other blood group systems have been discovered. All are assigned numbers and conventional abbreviations by the International Society for Blood Transfusion (ISBT). At the end of 2025, the ISBT listed 45 blood groups. There is insufficient space in this chapter to discuss all blood groups. Attention will be concentrated on the ABO and Rh systems; some others of biological interest and clinical significance will be discussed more briefly.

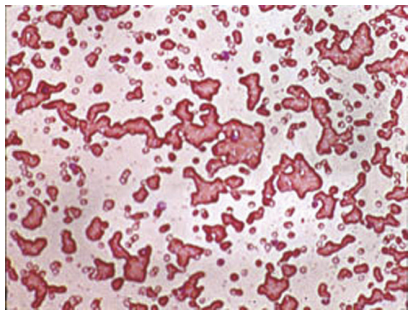


Figure 6.8 Photomicrograph showing erythrocytes agglutinated by antierythrocyte antibodies.

Margin Note 6.1
Landsteiner s Nobel Prize

The discovery of the ABO blood group system led to Karl Landsteiner receiving the Nobel Prize in Physiology or Medicine in 1930.

6.4 THE ABO BLOOD GROUP SYSTEM (ISBT 001)

The ABO blood group system classifies people into one of four major types, A, B, AB and O, according to the different types of antigen(s) present on the surfaces of their erythrocytes. The frequencies of these blood groups vary between different populations (*Table 6.2*).

Antigens of the ABO system

The antigens that determine ABO blood groups are glycans or oligosaccharides carried on cell surface glycolipids and glycoproteins that protrude from the erythrocyte plasma membrane (see *Figure 6.9*).

		Frequency (%)			
Phenotype	**Antigen**	**White**	**Indian**	**Black**	**Overall in UK**
A	A	44	24	27	43
B	B	9	38	19	9
AB	A and B	3	15	3	3
O	H	44	22	49	45

Table 6.2 Frequency (as %) of A, B, AB and O blood groups in different ethnic groups.

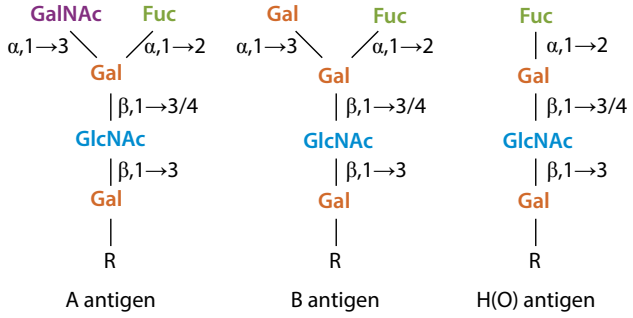


Figure 6.9 Structures of the ABO blood group glycans/oligosaccharides.

The gene *FUT1* (formerly *H*), located on chromosome 19, encodes a fucosyltransferase (FUT1), which adds an L-fucose residue to the precursor glycan forming the H antigen (*Figures 6.9* and *6.10*). In blood group O individuals, the H antigen is found on their erythrocytes and also on a variety of other cells. People of blood group A and AB possess the *A* gene on chromosome 9, which encodes an *N*-acetylglucosaminetransferase that adds an *N*-acetylgalactosamine residue to the terminal galactose of the H antigen to form the A antigen. This antigen is found on the erythrocytes of individuals of blood groups A and

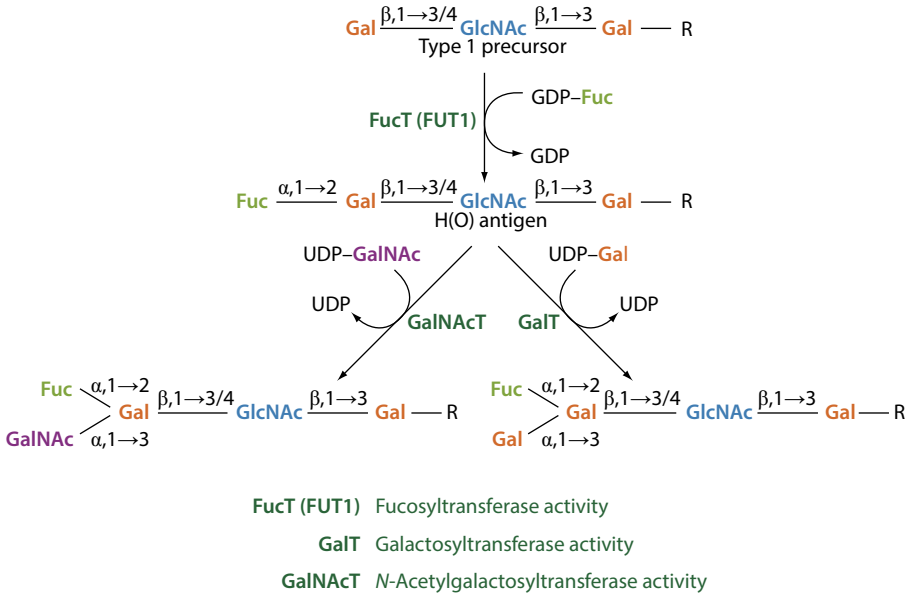


Figure 6.10 The biosyntheses of the ABO blood group antigen oligosaccharides/glycans.

AB. The *B* gene, which is allelic to the *A* gene, encodes a galactosyltransferase that adds a galactose residue to the galactose of the H antigen to give the B antigen (*Figures 6.9* and *6.10*). This antigen is found on the erythrocytes of blood groups B and AB individuals. The *O* gene is also allelic to the *A* and *B* genes, but does not appear to encode a protein, that is, it is a silent gene. The *A*, *B* and *O* genes are inherited in a Mendelian fashion (*Chapter 15*), and A and B are codominant. The genotypes that determine the different phenotypes in the ABO system are shown in *Table 6.3*.

Genotype	Comment	Blood group (phenotype)
AA	The gene encoding the 'A' glycosyltransferase is present on both copies of chromosome 9	A
AO	The gene encoding the 'A' glycosyltransferase is present only on a single copy of chromosome 9	A
BB	The gene encoding the 'B' glycosyltransferase is present on each copy of chromosome 9	B
BO	The gene encoding the 'B' glycosyltransferase is present on only a single copy of chromosome 9	B
AB	One copy of chromosome 9 has the gene encoding the 'A' enzyme while the other has the gene encoding the 'B' enzyme	AB
OO	Neither the 'A' nor the 'B' gene is present on either chromosome 9	O

Table 6.3 Genotypes and phenotypes of the ABO system. ⏎

Approximately 75% of the White population attach A, B and H antigens (depending on their blood group) to soluble proteins and secrete them into saliva and other body fluids and are known as **secretors**. Secretor status is dependent on the expression of a fucosyltransferase (FUT2), which is encoded by the gene *FUT2* (previously *SECRETOR*, *Se*) on chromosome 19. In the past, the presence or absence of secreted ABH antigens was used for identification in forensic science. The Lewis blood system (see later) is formed, in part, by the activity of FUT2.

Antibodies of the ABO blood group system

Transfusions of ABO incompatible bloods cause clinical problems because individuals possess antibodies in their plasma to the complementary antigens. Thus, people of blood group A have antibodies to the B antigen, those with blood group B have antibodies to the A antigen, blood group AB individuals have neither antibody and those of blood group O possess both (*Table 6.4*). These antibodies will bind to complementary antigens leading to agglutination of the erythrocytes. The pattern of agglutination when serum and erythrocytes of different ABO types are mixed *in vitro* is shown in *Table 6.5*.

Blood group	Antibodies in plasma
A	Anti-B
B	Anti-A
AB	Neither anti-A nor anti-B
O	Anti-A and anti-B

Table 6.4 Antibodies of the ABO system. ⏎

Antibodies directed against the ABO antigens generally belong to the immunoglobulin M (IgM) class, although immunoglobulin G (IgG) antibodies may also be found (*Chapter 4*). Immunoglobulin M antibodies are pentameric molecules (see *Figure 4.26*), which are efficient agglutinators of cells and activators of complement. These antibodies, which are sometimes called **isohaemagglutinins** because they agglutinate erythrocytes, are also termed 'natural antibodies' because they are present in people who have never been 'immunised' with blood of an inappropriate group. Of course, people do not

Cells	Serum			
	A	B	AB	O
A	Negative	Positive	Negative	Positive
B	Positive	Negative	Negative	Positive
AB	Positive	Positive	Negative	Positive
O	Negative	Negative	Negative	Negative

Table 6.5 Patterns of haemagglutination when sera and cells of different ABO types are mixed. ⏎

make antibodies to antigens that they themselves express, because of immunological tolerance (*Chapter 5*). Indeed, blood group antigens are also found on some common bacteria and the antibodies produced following infection can cross-react with blood group antigens.

Consequences of ABO incompatible transfusions

An example of an incompatible blood transfusion would be when the *recipient* of blood group A is given blood from a *donor* who is blood group B. In such a case, the anti-B antibodies in the plasma of the recipient bind to the donated erythrocytes and activate complement. This results in the simultaneous lysis of billions of erythrocytes in the donated blood. At the same time, the limited amount of anti-A antibodies in the donated blood bind to recipient erythrocytes, and these also haemolyse. The lysis of so many erythrocytes, called **acute intravascular haemolysis**, releases so much haemoglobin that acute renal failure and shock result. This has serious clinical consequences and, indeed, is fatal in approximately 10% of cases. In addition to the problems caused by the release of haemoglobin, the fragments of erythrocyte membrane produced may also initiate the blood clotting systems, leading to disseminated intravascular coagulation (DIC).

The consequences of an incompatible transfusion may vary according to the blood group involved. For example, antibodies to some of the other blood group systems may result in a delayed, extravascular haemolysis. Slower destruction of the donated cells may lead to a decreasing haemoglobin concentration, with the patient suffering a fever and general malaise.

The administration of an incompatible donation is most often due to errors post-donation and is rarely due to mismatching of bloods.

6.5 THE RHESUS BLOOD GROUP SYSTEM (ISBT 004)

The Rh blood group system divides people into Rh-positive (Rh+) and Rh-negative (Rh–) types depending on whether or not their erythrocytes carry the Rh antigen. Landsteiner and Alexander Wiener (1907–1976) discovered this system in 1940. They showed that antisera raised in guinea pigs against erythrocytes from Rhesus monkeys reacted with 85% of White population blood donors in New York. The Rh system of blood group antigens is often described as if it is a single antigen. However, it consists of a complex series of antigens, which are specified by two genes: *RHD* and *RHCE*. The former encodes the RhD protein, which expresses the D antigen, while the latter encodes RhCcEe proteins that carry the C or c antigen together with the E or e antigen. At one time it was thought that another antigen, termed the 'd' antigen, was present when the D antigen was absent. It is now recognised that the d antigen does not exist. However, the term is still used to indicate the D-negative phenotype.

An individual person inherits a set of genes from each parent, where the symbol '*d*' is used to express a lack of *RHD*. Under the Fisher system of nomenclature, each Rh haplotype is assigned a code. The possible haplotypes (haploid genotypes) are shown in *Table 6.6*. If an individual has the D antigen, they are said to be RhD positive. Thus, amongst the most common genotypes, the only individuals who are RhD negative are those with the genotype dce/dce (or rr using the Fisher code). Such genotypes are present in 15% of the White population. Individuals who are RhD negative do not normally have antibodies to the D antigen. However, they can become sensitised if transfused with blood from an Rh-positive individual. For example, a person with the genotype *dce*/*dce* who is

Haplotype	Fisher system code
DCe	R1
DcE	R2
DCE	Rz
Dce	Ro
dCe	r′
dcE	r″
dCE	r^y
Dce	r

Table 6.6 Haplotypes of the Rhesus blood group system.

transfused with blood from a *DCE/DCE* donor may make antibodies to the C, D and E antigens, although anti-D antibodies are the most common. *Table 6.7* summarises the Rh group genotypes and the Rh status in the White population.

Genotype	Frequency in the White population (%)	Rh status
DCe/dce	33	Positive
DcE/dce	11	Positive
Dce/dce	2	Positive
DCe/DCe	18	Positive
DcE/DcE	2	Positive
DCe/DcE	14	Positive
DCe/Dce	2	Positive
dce/dce	15	Negative

Table 6.7 The most common genotypes of the Rhesus (Rh) blood group system (in the White population).

6.6 OTHER BLOOD GROUP SYSTEMS

Many blood groups have been identified additional to the ABO and Rh systems. The following examples are described to give some flavour of their diversity:

- Kell
- Lewis
- Duffy
- Kidd

The Kell blood group system (ISBT 006, symbol KEL) is highly polymorphic and comprises more than 25 antigens, which are expressed on a metalloendopeptidase glycoprotein of the erythrocyte membrane. The antigen KEL (formerly Kell) and several others are highly immunogenic and IgM or IgG antibodies to them are common in transfused patients. Indeed, after the ABO and Rh antigens, those of the KEL system are the most likely to cause clinical problems following transfusions. Similarly, antibodies to KEL antigens can also result in haemolytic disease of the newborn (HDN; *Section 6.9*), although specific antibodies are rare in transfused patients.

The primary antigens of the Lewis blood group system (ISBT 007, symbol Le) present on erythrocytes are Le^a and Le^b (see also *Chapter 5* for other material relating to Lewis antigens). However, these antigens are not integral parts of the membrane but soluble

plasma proteins that become reversibly adsorbed onto erythrocyte membranes. The amounts of bound antigens therefore vary, although the erythrocytes of children two years old and above have approximately adult amounts.

The Le^a and Le^b antigens are the products of the fucosyltransferase FUT3. If the fucose residue is attached to the precursor of H (known as type 1 precursor oligosaccharide), then the Le^a antigen is produced (*Figure 6.11*). If, however, it catalyses the addition of a fucose residue to the H antigen formed by the action of FUT2 (see *Section 6.6*) then it forms the Le^b antigen. The activity of FUT3 can also convert the A and B antigens to the ALe^b to BLe^b antigens, respectively. Approximately 72% of white populations are Le(a–b+), that is, they lack the Le^a but have the Le^b; 22% are Le(a+b–); and 6% lack both antigens.

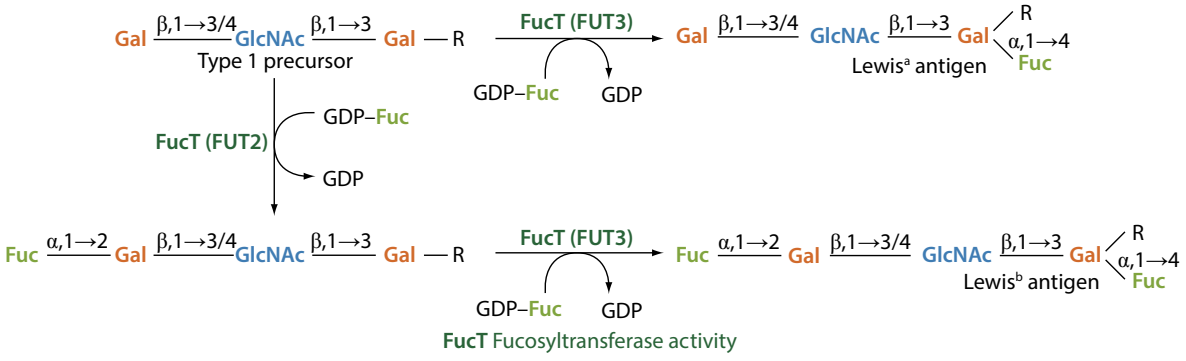


Figure 6.11 The roles of fucosyltransferases in the formation of Lewis^a and Lewis^b antigens.

Antibodies to the Le antigens are usually of the IgM class and, as such, do not cause HDN (*Section 6.9*) since they do not cross the placenta and damage the fetus.

The Duffy system of blood group antigens (ISBT 008, symbol FY) comprises six antigens, of which Fy^a and Fy^b are the most significant in transfusion reactions. These antigens are found on the atypical chemokine receptor 1 (ACKR1), a glycoprotein known as FY or CD234 (*Chapter 4*) expressed on the plasma membrane of erythrocytes and a wide variety of other cell types. This receptor binds both CXC and CC chemokines (*Chapter 4*) and thus has roles in inflammatory responses. Interestingly, it is also the receptor for malarial parasites (*Chapter 2*). Hence, an FY(a–b–) individual, who does not express these antigens, has a selective advantage in malarial areas. Indeed, 68% of Black ethnic groups of African descent are of this phenotype, which is rare in the White population, and so lack binding sites for *Plasmodium vivax*. Antibodies to Duffy antigens belong to the IgG class and may cause HDN (*Section 6.9*).

The Kidd antigens (ISBT 009, symbol JK) are expressed on a plasma membrane glycoprotein that is a urea transporter, encoded by *SLC14A1*, and expressed on erythrocytes and in the kidney. The Jk^a and Jk^b antigens result from the expression of a codominant pair of genes. Examples of ethnic frequencies of the four genotypes are given in *Table 6.8*. The rare Jk(a–b–) phenotype is not associated with any obvious clinical syndrome except for a reduced concentration of urea in the urine, and erythrocyte plasma membranes exhibit a reduced permeability to urea. Antibodies to Jk^a and Jk^b belong to the IgM or the IgG class and can cause a mild form of HDN (*Section 6.9*).

Margin Note 6.2 Glycosyltransferase specificities

Glycosyltransferases are highly specific for their nucleoside sugar donor substrate, the glycosidic bond they form and the acceptor substrate to which they attach the sugar residue. Given it can transfer its fucose residue to one of several substrates, the α,1→4 fucosyltransferase, FUC3, is an exception to this rule.

Genotype	Proportion (%)	
	Black	White
Jk(a+b+)	34	50
Jk(a+b–)	57	27
Jk(a–b+)	9	23
Jk(a–b–)	Rare	Rare

Table 6.8 Distribution of Kidd genotypes in Black and White groups.

6.7 LABORATORY DETERMINATION OF BLOOD GROUPS

Traditional methods to determine blood groups are based on **haemagglutination**, that is, the agglutination of erythrocytes, which occurs when the antibodies cross-link them, forming visible aggregates. Haemagglutination can be carried out on glass microscopy slides or in microtitre plates in which agglutination patterns are easily distinguished (see *Figure 6.8*).

The extent of haemagglutination depends on the temperature, pH and ionic strength of the medium. Agglutination is favoured in low ionic strength saline (LISS). Erythrocytes have a net electronegative charge and repulsive forces normally keep them about 20 nm apart. When antibodies bind to the erythrocyte, the reduced surface charge allows the cells to agglutinate. This is most effectively achieved with IgM antibodies, which can cause direct agglutination of erythrocytes. To obtain a direct agglutination with IgG antibodies, it is usually necessary to include bovine serum albumin in the medium, which also masks the charges on the erythrocytes and allows them to come closer together. Another method of reducing the negative charge is to use proteolytic enzymes to remove surface proteins that carry the charges. The enzyme can be added to the erythrocytes prior to the addition of the antibody, or all the components can be added together. Polycationic polymers such as polybrene will also reduce the negative charge on erythrocytes.

In *forward* typing, blood groups, such as in the ABO system, are identified by adding the individual's blood to known antibodies. In contrast, *reverse* typing involves adding the plasma or serum from the individual to known antigens. In either case the pattern of haemagglutination will enable the blood group to be determined. Performing both forward and reverse grouping acts as an internal check for accuracy. When different results are obtained, termed *discrepancies*, further tests are necessary to identify the blood group. Most discrepancies arise from technical errors or mislabelling of blood samples.

Clinical blood grouping or typing is largely done by automated systems using monoclonal antibodies. Recent decades have seen increasing use of the column or gel method to determine blood groups. Individual tubes, each containing a minigel in which monoclonal typing antibodies are distributed, are set in plastic 'cards'. Forward and reverse typing are performed simultaneously in different columns. The appropriate samples are added to their columns, time is allowed for agglutination to occur and the cards are centrifuged. Where agglutination has occurred, the agglutinates remain on top of the gel, whereas nonagglutinated cells settle through the gel to the bottom (*Figure 6.12*).

Margin Note 6.3
Discrepancies in ages

Reverse grouping is usually not performed for babies less than six months old. At this age, babies have very low concentrations of ABO antibodies (*Chapter 5*) and therefore present very weak reverse typing reactions. Given that, with increasing age, the immune system, like all body systems, loses efficiency and antibody production decreases, reverse blood grouping also diminishes as an effective check in the elderly.

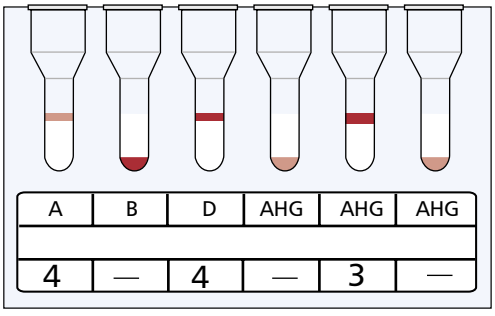


Figure 6.12 The column or gel method of determining blood groups. Agglutinated cells do not penetrate the gel. The blood group indicated here is A Rh+, as indicated by the first three tubes, where agglutination has occurred in tubes 1 and 3. AHG, anti-human globulin.

Some automated analysers use solid phase assay methods. Antibodies or antigens are fixed to the surfaces of microplate wells. The test samples are added to the wells and the plate centrifuged and washed. The resulting patterns of adhesion indicate the blood group. For example, in a forward test, the erythrocytes would adhere to the well with the complementary antibody.

Identifying blood groups by DNA-based genotyping is attracting interest as an alternative to the traditional serological antibody–antigen methods described above. Genotyping relies on molecular biological methods, such as the polymerase chain reaction (PCR; see *Chapter 3* and *Figure 6.27*). Such methods are particularly valuable when it is not possible or is technically difficult to obtain blood samples, such as in prenatal fetal diagnosis of HDN, but chorionic villus sampling can supply the required DNA. Difficulties in traditional blood grouping are also encountered when a patient has received large volumes of transfused blood and serology may well reflect the donor's status rather than that of the recipient.

6.8 COMPLEMENT AND BLOOD TRANSFUSIONS

The roles of complement are described in *Chapter 4*. However, given that it amplifies the actions of antibodies, complement can cause some of the problems associated with transfusion reactions. For example, complement causes lysis of *sensitised* erythrocytes, that is, those coated with antierythrocyte antibodies. Such antibodies may be present due to transfusion reactions or HDN (*Section 6.9*). The classical complement pathway is initiated when IgG or IgM binds to epitopes on the erythrocyte plasma membrane; usually IgM is the case with anti-A or anti-B, but IgG is the normal antibody to Rh antigens (*Figure 6.13*). Given IgM is the more efficient activator of complement, its presence is more of a concern. The binding of antibodies to these epitopes allows C1q to bind to the Fc regions of multiple antibodies and so initiates complement reactions. This results in the formation of membrane attack complexes (*Figure 6.14*), which allow Na^+ to enter the erythrocytes followed by water, causing their lysis. This is of greatest concern in transfusion reactions when antibodies in recipient blood bind to donated erythrocytes. However, the presence of antibodies in the donated blood can initiate an attack on recipient erythrocytes and this may also be of clinical significance.

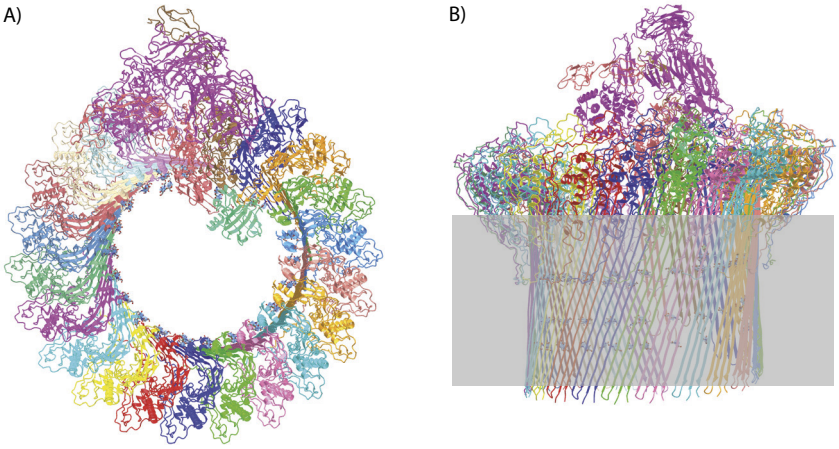


Figure 6.14 Molecular model showing (A) the cytoplasmic surface view and (B) the side view of a membrane attack complex. The position of the membrane lipid bilayer is indicated by the β strands in the side view. The grey shaded area indicates the position of the membrane. PDB file 6HO3. See also *Figure 4.11*.

Sensitised erythrocytes can also initiate their own destruction by complement-induced phagocytic cells that have receptors for C3b and other complement proteins on their surfaces (*Table 6.9*).

Receptor	Distribution on cells involved in clearing sensitised erythrocytes	Protein recognised by receptor
CR1	Erythrocytes, neutrophils, eosinophils, monocytes, macrophages	C3b, C4b, C3bi[a]
CR3	Neutrophils, large granular lymphocytes, macrophages	C3bi[a]
CR4	Neutrophils, monocytes, macrophages	C3bi[a]

[a] An inactive form of C3b.

Table 6.9 Complement receptors.

Receptors for C3b are found on all the major types of phagocytic cells, although macrophages in the spleen and liver are mainly responsible for removing erythrocytes coated in antigen–antibody complexes bound to C3b. Also, these organs have macrophages with plasma membrane surface receptors that are able to bind to Fc regions of antibodies and initiate the destruction of sensitised erythrocytes.

Other complement proteins can trigger inflammatory reactions. For example, C3a, C4a and C5a cause basophils and mast cells to degranulate. The mediators released stimulate inflammation (*Chapter 4*), which may have consequences in a patient who has antierythrocyte antibodies. In addition, both C3a and C5a are chemotactic factors for neutrophils and promote a build-up of these cells, which may itself lead to clinical problems.

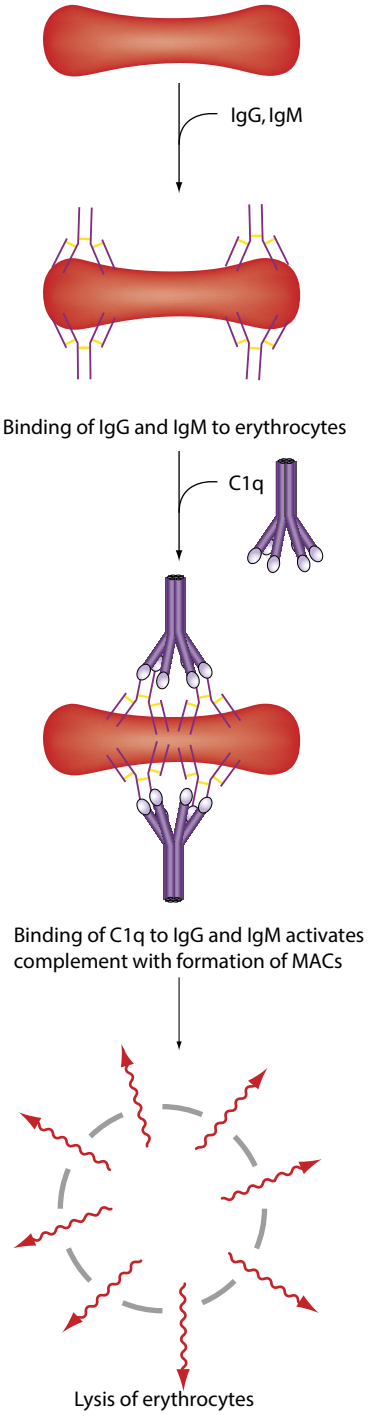

IgG, IgM
Binding of IgG and IgM to erythrocytes
C1q
Binding of C1q to IgG and IgM activates complement with formation of MACs
Lysis of erythrocytes

Figure 6.13 Complement activation and haemolysis of sensitised erythrocytes. Ig, immunoglobulin; MAC, membrane attack complex.

In vivo, the presence of regulatory proteins means that many antibodies with the potential to promote the lysis of erythrocytes are prevented from doing so. For example, one of the regulatory steps of the alternative pathway for complement activation (*Chapter 4*) is the binding of C3b to factor H. In this complex, C3b is inactivated by factor-I (see *Table 4.2*) that converts it to the inactive C3bi, which can be subsequently degraded to the fragments C3dg and C3d.

It is, naturally, essential that any haemolytic antibodies in donor or recipient blood be detected before a transfusion. The presence of relevant antibodies may be detected in samples of serum by haemolysis *in vitro* using the complement activity in the serum or they may be detected by examining the surface of erythrocytes for activated complement proteins. In practice, it is easier to investigate erythrocytes for the presence of complement proteins than to look for antibodies because a small deposit of antibody can result in large amounts of complement activity. *In vitro*, the sudden clearing of the opaque suspension of erythrocytes upon the addition of a sample of serum is indicative of their lysis by complement.

Given their short half-lives, the presence of complement proteins diminishes with storage; therefore, it is recommended that samples of sera should be stored at –20°C to retain activity if they are to be used for haemolysis determination. Further, some anticoagulants, such as ethylenediaminetetraacetate (EDTA), inhibit complement, which may be significant if plasma rather than serum is available. Other sera may have 'anti-complement' activity due to the presence of a denatured form of complement known as complementoid.

6.9 HAZARDS OF TRANSFUSION

Blood transfusions are among the safest of all clinical interventions. However, medical problems can occur, some of which are described below.

Haemolytic disease of the newborn

Haemolytic disease of the newborn is a serious disease characterised by anaemia, splenomegaly, hepatomegaly and oedema. The most common cause of HDN is antibodies against the RhD antigen, although it can also be caused by antibodies to other blood groups, for example to A or B antigens, if the antibodies are of the IgG class. In addition, antibodies to the Rhc and to the Kell blood group antigens (*Section 6.6*) may also be involved.

The condition results from the transfer of maternal antierythrocyte antibodies across the placenta. It may also arise in the offspring of women who are RhD negative. Such women may become sensitised to RhD antigens when they give birth to an RhD-positive child. At birth, some of the baby's blood can enter the maternal circulation and the mother will respond by making anti-D antibodies. This does not have any clinical consequence for that baby but may cause problems during a subsequent pregnancy if the fetus is RhD positive. Antibodies to Rh antigens are clinically significant during pregnancy because, being of the IgG class, they are able to cross the placenta and bind to fetal erythrocytes. Although Rh antibodies are IgG, they do not appear to activate complement. Instead, monocytes and macrophages have receptors for the Fc portion of IgG molecules and can therefore bind to the antibody-coated erythrocytes. This phenomenon, called **immune adherence**, occurs when blood circulates through the spleen and liver. The uptake and subsequent destruction of antibody-coated erythrocytes in these organs is referred to as **extravascular haemolysis**.

The consequences for the second and/or subsequent children depend on the extent of maternal sensitisation and, hence, the amount of circulating antibody. Destruction of fetal erythrocytes *in utero* may lead to fetal anaemia and hyperbilirubinaemia, an excess of bilirubin in the blood (*Chapter 13*). The concentrations of serum bilirubin indicate the degree of haemolysis. Babies born with significant levels of bilirubin will suffer **kernicterus** or brain damage due to an accumulation of lipid soluble bilirubin in the brain. The baby will need phototherapy, which helps to break down the bilirubin (*Chapters 11* and *13*). Severe anaemia may result in the death of the fetus due to heart failure or hydrops fetalis, an extreme oedema affecting the entire body of the fetus, which may lead to a spontaneous abortion. A longer term child may be stillborn. The considerably enlarged spleen and liver in a child born alive with HDN is associated with erythrocyte destruction in these organs. The baby may also have a facial rash indicative of haemorrhages due to impaired platelet functions.

All pregnant women in the UK and other developed countries have their ABO and Rh status checked at their initial hospital booking, which is usually at 12–16 weeks of pregnancy. They are also checked for anti-D, or anti-Kell, antibodies (*Section 6.6*) and, if these

Margin Note 6.4
HDN and the ABO system

Haemolytic disease of the newborn does not occur within the ABO system if the anti-A and anti-B antibodies are IgM because this class cannot cross the placenta. In addition, a woman who is blood group A, RhD negative, is unlikely to become sensitised to a fetus of blood group B, RhD positive, because, when fetal erythrocytes enter her circulation following birth, her anti-B antibodies will destroy the fetal erythrocytes before sensitisation to RhD can occur.

are present, their concentrations will be monitored regularly throughout the second trimester of pregnancy. If the concentrations of clinically significant antibodies start to increase, clinical intervention may be necessary. Concentrations of anti-D of less than 4 international units (IU)/mL are unlikely to cause HDN, 4–15 IU/mL have a moderate risk of HDN and greater than 15 IU/mL are associated with a high risk of HDN.

HDN prophylaxis

Since the 1970s, a prophylactic treatment to prevent Rh sensitisation has been available that has greatly reduced the incidence of HDN from RhD sensitisation. The treatment involves intramuscular injections of at least 500 IU of anti-D immunoglobulin within 72 hours of the birth of an RhD-positive baby. These antibodies bind to any erythrocytes from the baby that have entered the maternal circulation and destroy them, preventing the mother from making antibodies.

In 2002, the National Institute for Health and Care Excellence (NICE) in the UK recommended that antenatal anti-D prophylaxis should be routinely offered to any pregnant Rh-negative woman who has not made anti-D antibodies to prevent sensitisation and predelivery, caused by, for example, a placental bleed.

Direct antiglobulin and Kleihauer testing

The direct antiglobulin test (DAT), formerly known as the direct Coombs test, is undertaken to determine if maternal antibodies are present on the baby's erythrocytes. The test uses antihuman globulin (AHG) to agglutinate erythrocytes coated with nonagglutinating erythrocyte-specific IgG. The AHG can be used to detect erythrocytes already coated with antierythrocyte IgG molecules in the DAT or can be used on cells that have been incubated with antibody *in vitro*. If they are present, then a sample of erythrocytes from the baby will be agglutinated by AHG (*Figure 6.15*).

The Kleihauer test uses the Kleihauer–Betke stain to assess the volume of fetal blood that has entered the maternal circulation. In most cases this volume will be <4 mL, and 500 IU of anti-D is sufficient to remove the erythrocytes from these volumes. However, for less than 1% of mothers, the volume of fetal blood is larger and consequently more than this quantity of anti-D is required. The Kleihauer test is carried out on a sample of maternal blood two hours after delivery. The principle of the test is that the haemoglobin in adult erythrocytes can be eluted with acid, whereas that in fetal ones is resistant to acid elution. A smear of maternal blood is placed in a solution of haematoxylin and hydrochloric acid at pH 1.5 and then counterstained with eosin. The maternal erythrocytes appear as pale 'ghosts' but eosin stains fetal ones pink and leukocytes blue (*Figure 6.16*). The ratio of fetal to maternal cells is an indicator of the volume of blood that has entered the circulation. This test can also be performed during pregnancy if a placental bleed is suspected.

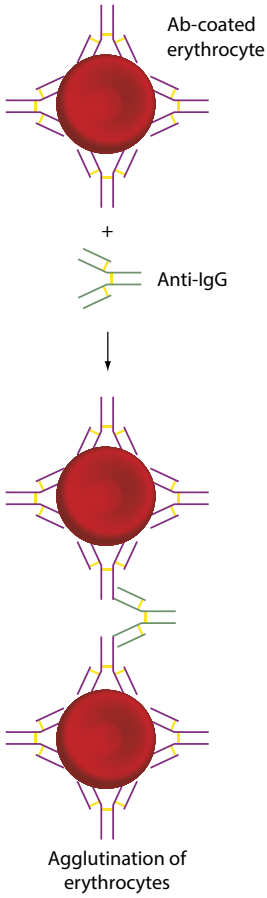


Figure 6.15 Schematic illustrating the direct antiglobulin test.

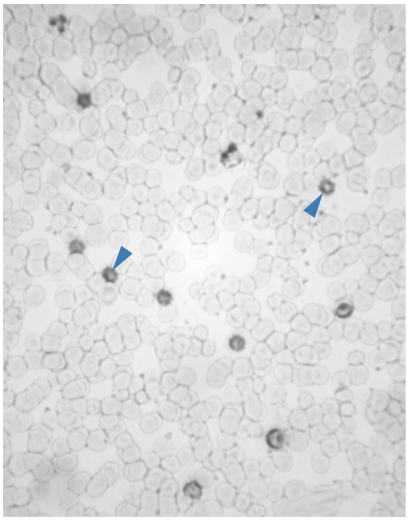


Figure 6.16 Photomicrograph showing the results of a Kleihauer test. The darker-staining erythrocytes are fetal cells.

Haemolytic transfusion reaction

A haemolytic transfusion reaction (HTR) is a hazard of transfusion if preexisting antibodies (see *Box 6.4*) are present in the recipient. This may result in acute intravascular haemolysis, as in ABO incompatibility, or in delayed extravascular haemolysis, as with several of the other blood group systems. Acute intravascular haemolysis has serious clinical consequences and, indeed, may be fatal. With delayed extravascular haemolysis, the patient may suffer fever and general malaise as the donated erythrocytes are destroyed and the haemoglobin concentrations decrease. To prevent either occurrence, the recipient should be screened before the transfusion for the presence of antierythrocyte antibodies.

The transfusion of leukocytes may also lead to adverse reactions, such as **nonhaemolytic febrile transfusion reactions**. Patients exhibit flushes, fever, rigours and hypotension. These may be caused by the reaction between antibodies and leukocyte antigens in the recipient, resulting in the lysis of donated leukocytes and release of cytokines from them. In addition, activated complement proteins cause the release of histamine from basophils, triggering inflammatory reactions. As all blood is leukodepleted before transfusion, such reactions are now rare.

Transfusion-related acute lung injury

Transfusion-related acute lung injury (TRALI) is a life-threatening complication of transfusion. It presents as rapidly progressing and severe respiratory failure with diffuse damage to alveolar cells and the filling of alveolar spaces with fluid. Histological examinations reveal an infiltration of the alveoli with neutrophils and monocytes, indicative of an acute inflammatory reaction.

The symptoms of TRALI include dyspnoea, cyanosis, hypotension, fever and pulmonary oedema, which usually occur within six hours of transfusion. The condition is thought to arise from the interaction between leukocytes and specific antileukocyte antibodies, usually in the donor plasma, although occasionally in that of the recipient. The most likely antibodies are those to the MHC antigens (*Section 6.12* and *Chapter 4*), which are most common in women who have had multiple pregnancies and in males and females who have received multiple transfusions. In addition, antibodies to neutrophil antigens have also been implicated. In 2004, 23 suspected cases were reported in the UK. Of these, 13 were highly likely to be TRALI and were linked to fresh frozen plasma in six cases, platelets in four, erythrocytes in two and whole blood in one. The incidence of TRALI has since decreased in the UK, due to the processing of fresh frozen plasma from untransfused male donors who have tested negative for antileukocyte antibodies.

Iron overload

Patients who receive many transfusions over long periods may develop iron overload. The excess of iron, for which there is no excretory route (*Chapter 13*), from the transfused blood may cause tissue damage, especially to the liver, heart and endocrine glands. Signs of iron overload can be detected after 10–20 transfusions. The condition may be fatal and patients should be treated with chelating agents, such as deferrioxamine mesylate (*Figure 6.17*), to remove unwanted iron.

Alloimmunisation

Patients who receive regular erythrocyte transfusions may become immunised to other blood group antigens present on the ABO compatible cells. This may have consequences for future transfusions. If patients receive whole blood, rather than leukodepleted blood, they may become immunised to the MHC antigens on the foreign leukocytes. This could become clinically significant if in the future they require a tissue/organ transplant (*Section 6.11*). However, the use of leukodepleted blood, in which the leukocytes have been removed prior to transfusion, prevents the immunisation of recipients to these antigens.

Graft versus host disease

Graft versus host disease is a potentially fatal condition associated with transfusions of whole blood or blood products, such as packed erythrocytes or platelets that contain residual lymphocytes. It usually results from the transfusion of leukocytes into an immunodeficient patient or when blood is transfused into premature babies or neonates. Small lymphocytes present in the donated blood recognise the antigens of the recipient as foreign and mount an immune response against them. The donor lymphocytes proliferate in the patient and attack tissues, causing diarrhoea, enlargement of the liver and spleen, and extensive skin rashes. Acute GVHD can be fatal and for this reason products, such as packed erythrocytes, should be irradiated with X- or γ rays prior to their use, which prevents residual small lymphocytes from reacting to the host antigens. Graft versus host reaction can also be a consequence of transplanting solid tissues or organs or haematopoietic stem cell transplantations (HSCTs; see *Box 5.2* and *Section 6.14*). Frozen plasma is safe in this respect, since freezing destroys leukocytes. Transfusion-associated GVHD is not usually linked with acquired immunodeficiency syndrome (AIDS).

Infections

One potential hazard of transfusion is infection with viruses or microorganisms present in the donor. In the past, transfusions of blood and blood products, such as factor VIII needed by haemophiliacs (*Chapter 13*), have spread infections such as human immunodeficiency virus (HIV; *Chapters 2* and *3*) and hepatitis C (*Chapter 11*) to the patient. For this reason there is now extensive screening of blood donors (*Section 6.10*).

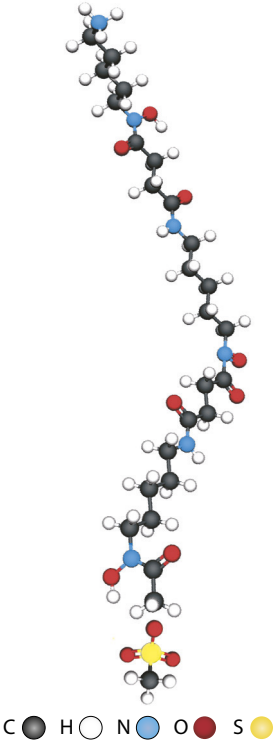


Figure 6.17 Computer drawn model of deferrioxamine mesylate.

6.10 SCREENING OF BLOOD DONORS

The transfusion of blood must pose minimal risk to patient and donor alike. Aside from the obvious need for blood to be carefully matched to avoid a possibly fatal transfusion reaction, it is essential that donors are carefully screened to avoid those who are ill, who may be harmed by giving blood or whose blood poses a health risk because, for example, it is contaminated with certain viruses even though the donor appears healthy.

In the UK, blood is taken from healthy donors aged 17–65 years weighing 50–158 kg and is voluntary and an unpaid activity. Potential donors who are excluded from

Margin Note 6.5
The Serious Hazards of Transfusion Group

In the UK, adverse reactions to transfusion are reported to the Serious Hazards of Transfusion (SHOT) group based at the Manchester Blood Centre. Here, the data are collected and an annual report is produced. This process enables trends to be recognised and recommendations on safe practices in transfusions to be made.

donation include individuals with HIV or hepatitis viral infections, as well as individuals who are at risk of becoming HIV and/or hepatitis virus positive, for example prostitutes, drug abusers who inject themselves with drugs, and individuals who have had sex with men or women infected with these diseases. *Table 6.10* lists some additional reasons for exclusion.

Category	Examples
Donor has history of clinical intervention	Had a local anaesthetic for dental treatment within the last seven days Had a miscarriage within the last 12 weeks or had a child within the last six months Had a major operation within the last six months Had received blood, platelets, plasma or any other blood products after 1 January 1980 Had an organ transplant
Donor is currently unwell	Is currently receiving medical treatment or under medical investigation Has a haemoglobin concentration below 135 g/L (male) or 125 g/L (female) Is currently taking aspirin or has taken antibiotics within the last seven days
Donor has recently been unwell/is unwell	Had chicken pox, cold sores, herpes simplex infection, measles, mumps, rubella or shingles within the last three weeks Had a cold, cough or sore throat in the last seven days Is known to be infected with HIV or hepatitis viruses
Donor is in a high-risk category for HIV and/or hepatitis viruses	Has injected recreational drugs or non-prescribed drugs including body-building and injectable tanning agents Had sex with a partner who is HIV positive, human T lymphotropic virus type 2 (HTLV) positive, a hepatitis B/hepatitis C carrier or syphilis positive Has received money or drugs for sex within the last three months

HIV, human immunodeficiency virus.

Table 6.10 Some of the reasons for exclusion from donating blood.

All donated blood is screened for a variety of infectious agents, as shown in *Table 6.11*. Some tests are mandatory while others are optional. Optional tests such as those for cytomegalovirus (CMV; *Figure 6.18*) are used when the blood is to be transfused into immunocompromised individuals.

Mandatory tests	Optional tests
Hepatitis B virus (HBV)	Cytomegalovirus (CMV)
Hepatitis C virus (HCV)	Malaria (*Chapter 3*)
Human immunodeficiency viruses (*Chapter 3*)	
Syphilis (*Chapter 3*)	

Table 6.11 Mandatory and optional tests for infectious agents in donated blood.

6.11 TRANSPLANTATION OF SOLID TISSUES AND ORGANS

The first successful kidney transplant was performed by Joseph Murray (1919–2012) in 1954, while the first human heart transplant, by Christiaan Barnard (1922–2001), was performed in 1967. The surgical transplantation of solid tissues or organs from one individual to another for therapeutic purposes is now a relatively routine medical procedure. For example, over 27 000 kidney and over 4500 heart transplants were performed in the USA in 2023.

Interest in transplantation mainly arose during the First and Second World Wars when badly burned combatants were treated using skin grafts from unrelated donors. Indeed, it was from experiments with skin grafting in rodents that the role of the immune system in the rejection of transplants was identified. Moreover, once this role was recognised, the search for immunosuppressive drugs became more focused.

Margin Note 6.6
Plasmapheresis and leukapheresis

Different components of the blood can be isolated while the donation is occurring by sending the blood through a cell separator. A catheter is first inserted into a vein of the donor. Blood entering the catheter is transferred to the cell separator where it is centrifuged at a speed necessary to collect the required component. The blood, minus the component that has been removed, is returned to the donor via a catheter inserted into a vein in the other arm. The term **apheresis** is a general term that covers the collection of specific blood components in this way (see also *Box 5.2*). Thus:

- plasmapheresis is the collection of plasma, with the return of the erythrocytes and leukocytes to the donor
- leukapheresis is the collection of leukocytes, with plasma and erythrocytes being returned
- plateletpheresis involves the collection of platelets, with all other components being returned

Margin Note 6.7
Blood donation in the USA

In the USA, blood donation is regulated by the United States Food and Drug Administration (FDA) and the Association for the Advancement of Blood and Biotherapies (AABB), which together, ensure the process of blood donation is safe.

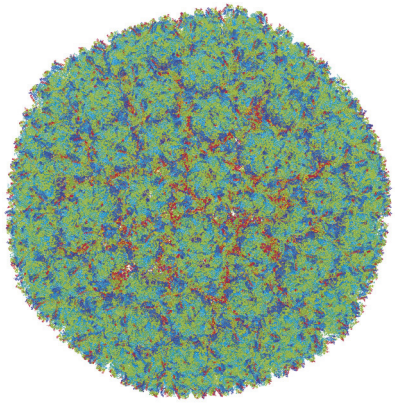


Figure 6.18 Model illustrating the capsid structure of the human cytomegalovirus. PDB file 6NHJ.

BOX 6.2 MEDIEVAL/RENAISSANCE NOSE JOBS

A somewhat unusual form of skin transplant was a type of rhinoplasty (nose reconstruction) practice in Renaissance Italy. Prior to these developments, noses damaged or lost in duels, warfare or diseases such as syphilis or to tumours that destroyed the nasal cartilage were reconstructed by partially excising a nose-shaped section of skin from the forehead or cheek. The partial excision left a flap of skin to maintain a blood supply to the graft. The flap was twisted, placed in position to cover the damaged area of the nose and sewn into place. Provided the graft grew in place and the patient did not succumb to infection, this procedure did provide a passable replacement nose. However, the patient was left with disfiguring scars on the forehead or cheek. Antonio Branca, operating in the 1400s, developed a new technique that left the face unscarred. He traced a template of a nose on the patient's arm to identify an appropriate area of tissue to form the graft. The skin was again partially excised, leaving a flap still attached to maintain its viability. This skin was then attached to the nose and the arm firmly bound to the head (*Figure 6.19*). After a suitable period, perhaps 10–14 days, when the graft had healed but was still attached to the host tissue, the lower part of the skin flap was cut to separate the arm, which no doubt required considerable physiotherapy, from the new nose. The new nose was then moulded into shape and slits cut for nostrils.

This procedure was later popularised by Gaspare Tagliacozzi (1545–1599), whose book, *De Curtorum Chirugia per Insitionem* (*On the Surgery of Mutilation by Grafting*), published many years later, described these procedures. However, after his death the surgical procedure fell into disuse.

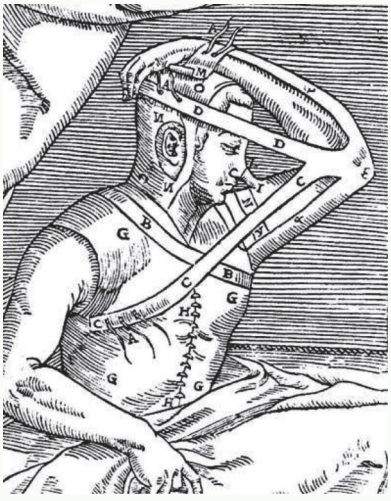


Figure 6.19 Rhinoplasty as practised by Branca in the 1400s.

Once technical problems associated with the transplantation of whole organs were overcome, the major problem associated with transplanting organs was, and remains, rejection. This is caused by the immune system of the recipient recognising the cells of the donated organ as foreign and mounting an immune response against them. Unless steps are taken to prevent it, transplant rejection is inevitable, unless the donor and the recipient are identical twins. However, organ transplantation is now remarkably advanced compared with initial attempts. Indeed, a range of transplants are routinely performed (*Table 6.12*). In some cases, combinations of organs, for example heart with lung and kidney with pancreas, are surgically transplanted together. These combined types of transplantation appear to induce fewer rejection problems than those encountered with single organ transplants.

Tissue/organ transplanted	Examples of clinical conditions
HSCT[a]	Immune deficiency; in treatment of leukaemia
Cornea	Types of blindness
Heart	Heart failure
Heart and lungs	Cystic fibrosis
Kidney	End-stage renal failure
Liver	Cirrhosis
Pancreas	Type 1 diabetes mellitus
Skin	Treatment of burns

[a] Haematopoietic stem cell transplantation (see *Box 5.2* and *Section 6.14*).

Table 6.12 Examples of tissues and organs transplanted.

Immunological rejection of transplants

The rejection of a transplant is due to the immune system of the recipient recognising the donated cells as foreign, that is, nonself. Thus, the greater the genetic disparity between

the donor and recipient, the greater the chances of rejection. The most common type of clinical transplant is an **allograft**, that is, a graft between two genetically nonidentical people. However, **isografts**, in which pieces of tissue are transplanted from one site to another on the same individual, are also routinely undertaken, for example in skin grafting to treat burns. Isografts are not rejected, since the donor and recipient tissues are, of course, genetically identical. Grafts may come from another species, termed **xenogeneic** grafts. More common examples of xenogeneic grafts involve the use of porcine and bovine heart valves to replace diseased human ones. In these cases, the valves are treated with glutaraldehyde to stiffen them and to mask their antigenic determinants. These are often referred to as **bioprosthetic** valves. In 1984, a baboon heart was transplanted into a terminally ill human baby who survived for 21 days. Interest has focused on the use of genetically engineered (GE) pigs as organ donors with some success. For example, baboons that received GE pig hearts have survived for over two years, but transplants of pig kidneys and hearts into humans have been rather less successful (*Box 6.3*).

Causes of graft rejection

The transplanted cells of an allograft carry histocompatibility antigens on their membranes, which are recognised as foreign by the cells of the recipient's immune system. Histocompatibility antigens are encoded by the MHC, which, in humans, is also known as the HLA complex (*Chapter 4*).

The MHC is a genetic region that encodes several classes of proteins, some of which are membrane proteins. Class I MHC proteins (MHC I) are present on the plasma membranes of all nucleated cells and are involved in the recognition of virus-infected cells by the precursors of cytotoxic T lymphocytes (T_C cells). Class II MHC proteins (MHC II) are found on the membranes of antigen-presenting cells (APCs), such as macrophages, and are involved in the recognition of foreign peptides on the surface of APCs by T_H lymphocytes. It is relatively minor differences between the primary structures (amino acid sequences) of the MHC molecules of the donor and those of the recipient that lead to the rejection of the transplanted tissue/organ. However, even when tissue typing shows the donor and recipient are congruent in their MHC genes, minor histocompatibility differences (MiHA) can lead to clinical problems. Minor histocompatibility proteins (antigens) are encoded by genes outside the MHC and are polymorphic in the population. Such antigens are found on the cell surfaces of donated tissues/organs and differences in the primary structures between donor and recipient can lead to GVHD-induced rejection of the transplant, although the frequency of rejection problems is less than with MHC antigens.

The rejection of an allograft may be:

- hyperacute
- acute
- chronic

BOX 6.3 GENETICALLY ENGINEERED PIGS AND XENOGENEIC TRANSPLANTATION ⏎

The chronic lack of organs available for transplant compared with the numbers and, indeed, the many deaths of patients needing them has promoted interest in xenogeneic transplants. Naturally, such transplants are subject to a greater risk of rejection owing to the large genetic disparity between the donor and the recipient.

There is much debate about the breeding of genetically modified pigs to provide organs for transplants into humans. Apart from the ethics of breeding animals purely to supply organs for humans, there are social concerns as to who will be chosen in clinical trials and about the enormous costs that will be associated with such therapies. In addition, there is also the risk of transferring an unknown, or indeed known, virus from pigs to humans, with possible clinical consequences. However, the use of pigs as organ suppliers offers several advantages: their physiology is not dissimilar to and their circulatory systems resemble that of humans; they are also easy to breed, producing large litters of piglets that grow rapidly to the point where their organs are comparable in size to those of patients.

In recent years, the development of genetically engineered (GE) pigs has been greatly assisted by gene editing (*Chapter 17*). Ten genetic modifications have been introduced (10-GE): four involved knocking out pig genes and six introduced human genes into the pig genome. Three of the porcine genes knocked out involved pig glycobiology, namely the α,1→3-galactosyltransferase, *N*-acetylneuraminic acid hydroxylase and the β,1→4-*N*-acetylgalactosaminyltransferase 2. The antigenic products of these genes promote graft rejection. This is notably the case in the product of the first gene, commonly referred to as the α-Gal epitope (*Figure 6.20*), which is widespread in animals but absent in humans and stimulates a hyperacute rejection if present on transplanted organs (see *Figure 6.21* and its associated text). The fourth pig gene knocked out was the porcine growth hormone receptor, which is necessary to prevent the donated organ growing too large in the recipient. The six human genes introduced included the complement inhibitory genes for CD46/complement membrane cofactor protein CD46 (*CD46*) and CD55/decay-accelerating factor for complement (*DAF*) to offer

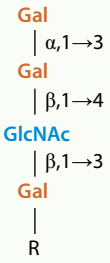


Figure 6.20 Structure of the pig α-Gal epitope. Compare its structure with that of the ABO blood group glycans in *Figure 6.9*.

protection against complement-mediated attack (*Chapter 4*). Two genes that encode thrombomodulin (*THBD*) and protein c receptor (*PROCR*) were incorporated to overcome coagulation side effects. The final two human genes were those for haem oxygenase 1 (*HMOX1*) and CD47 glycoprotein/integrin-associated protein (*CD47*) with the aim of reducing any inflammatory responses subsequent to organ transplantation. These pigs were developed with the expectation that such modifications would increase the compatibility of the transplanted organ(s) with the immune system of the patient.

In 2021, a clinical team from the University of Alabama transplanted kidneys from a 10-GE pig into a brain-dead human following permissions from relatives of the deceased and the Food and Drug Administration (FDA). The kidneys incorporated a thymic autograft from the pig under their capsules, forming a so-called *thymokidney*, as an approach to reducing the risk of host T cell mediated immune reactions (*Chapter 4*). The circulatory and respiratory activities of the cadaver were maintained by artificial ventilation. Over the three-day experiment, the kidneys functioned, producing a urinary output, although creatinine concentrations decreased. No bleeding episodes or signs of hyperacute rejection were observed. The value of this experimental study has, however, been criticised in some quarters, but the Alabama clinical team considered this approach an essential prerequisite to a 10-GE kidney transplant into a living patient.

In 2022, clinicians at the University of Maryland Medical Center transplanted the heart from a 10-GE pig into a 57-year-old male; this pioneering surgery was the result of more than 30 years of research. The patient was extremely weak and had relied on a heart–lung bypass machine. He was in end-stage heart failure and a transplant was regarded as the only possible therapy to save his life. However, he was considered unsuitable for a human transplant because of his weakened state and a possible history of poor compliance with prescribed medication. Permission for the experimental procedure was given by the FDA under its compassionate use provision. The surgical transplant of the heart was successful, and the patient was, naturally, subjected to aggressive antirejection therapy.

Electrocardiogram measurements showed the transplanted heart had longer PR intervals and increased QRS and QT periods (*Chapter 13*) than hearts involved in pig-to-pig transplants; in essence, the electrical signals were travelling slower than those in a healthy human heart. If such changes occurred in a human heart, they may indicate disease or abnormal rhythms (*Chapter 13*) but this does not appear to have been a problem. Unfortunately, the patient did develop infections that required several medical interventions. Following the transplant, extremely low titres of porcine cytomegalovirus (CMV) were detected in the heart, which is somewhat disconcerting given the donor pig was regarded as free of the virus. However, heart function was initially sufficient to improve the patient's condition. Sadly the patient suffered myocardial pathological trauma and an acute diastolic failure. The irreversible damage to the xenograft meant there was no option but to turn off life support systems 60 days after the transplant with the death of the patient.

Initial postmortem reports from the transplant centre reported a thickening and subsequent stiffening of the cardiac muscle, massive damage to capillaries with loss of fluid and the heart doubling its transplant weight. They concluded the patient died of heart failure and not from an acute rejection of the heart. Several possible factors have been proposed to account for the failure of the transplant. The patient was given intravenous immunoglobulin to combat an infection, which may have contained antipig antibodies or upset the balance of the antirejection drug regime. However, some clinicians have questioned these conclusions and suggested the damage to capillaries was indicative of organ rejection. All are agreed, however, that the extremely poor clinical state of the patient mitigated against any long-term survival. Transplants involving 10-GE pig organs are marked steps in xenotransplantation. It is hoped the findings of such experiments will be valuable learning experiences and inform future studies and perhaps lead to clinical trials.

Hyperacute rejections occur within minutes or hours of the transplanted graft becoming revascularised. Antibodies against graft antigens must already be present in the plasma of the recipient (*Box 6.4*), which bind to endothelial cells in the donated organ. This activates the subsequent infiltration of neutrophils and platelets that attack the donated organ and cause clots to form (*Figure 6.21*). The patient presents with extensive thrombosis and fever, and anuria if the graft involved is a kidney. Hyperacute rejection is irreversible: the only treatment is to remove the transplanted organ. Given that preexisting antibodies can cause a rapid rejection of a graft, it is essential to know which potential recipients have such antibodies. Hence, cross matching is performed, in which serum from the recipient is incubated with cells from the donor. If the donor cells are killed in the presence of recipient serum and complement, the transplant will not be undertaken and another potential recipient will be sought.

An **acute rejection** occurs within weeks of a transplant. Small lymphocytes recognise the transplanted cells as foreign and T lymphocytes infiltrate the graft. The presence of T lymphocytes and monocytes in the infiltrate is a strong indicator of cell-mediated immunity. Both T_H and T_C cells are involved in graft rejection. The T_C cells develop into cytotoxic T lymphocytes directed against the foreign histocompatibility antigens of the grafted cells. These are able to destroy cells of the grafts directly, or indirectly by producing

BOX 6.4 HOW CAN A PATIENT HAVE PREEXISTING ANTIBODIES TO A DONATED ORGAN?

It may be wondered why an organ recipient would have preexisting antibodies against the graft that can cause a hyperacute rejection. Such antibodies may be present for a number of reasons. For example, women who have had a number of pregnancies often have antibodies to the major histocompatibility complex (MHC) antigens on the fetus that were inherited from the father. Secondly, patients who have had a number of blood transfusions may become immunised to the MHC antigens on residual leukocytes present in the transfusion. Thirdly, patients who have had a previous transplant and rejected it will almost certainly have antibodies to any foreign MHC antigens that were present on that graft. Finally, antibodies to blood group antigens can also cause a hyperacute rejection if they are already present in the recipient. For example, blood group A, B and H antigens are present on endothelial cells that line blood vessels. Hence, if a recipient of blood group A is given, for example, a kidney from a person of blood group B, the anti-B antibodies in the recipient's plasma will attack the endothelial cells of the graft and activate complement leading to its destruction. For this reason, transplants are not carried out against major blood group barriers.

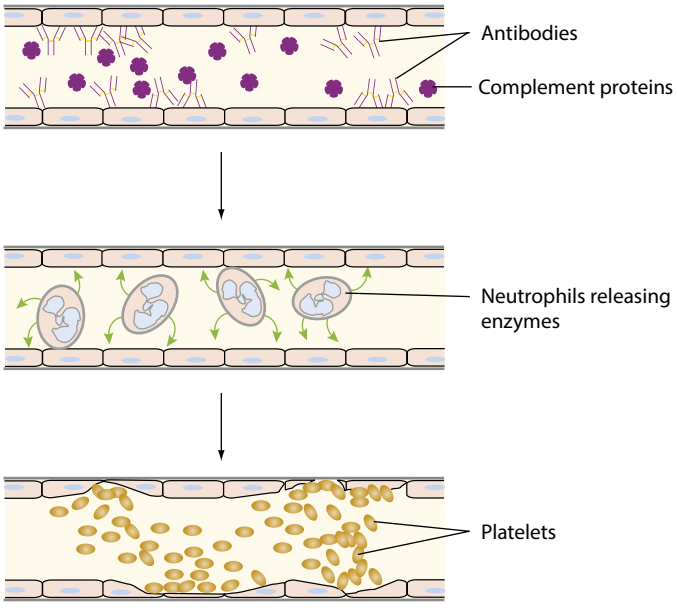


Figure 6.21 Schematic illustrating a hyperacute rejection of a transplanted organ.

cytokines that attract monocytes and macrophages. The T_H lymphocytes respond by producing cytokines that activate a variety of nonspecific cells to attack the graft.

Chronic rejections take place months or years after transplantation. They are brought about by combinations of humoral and cell-mediated mechanisms.

6.12 THE MHC OR HLA SYSTEM

The MHC or HLA system of genes is found on the short arm of chromosome 6 and encodes MHC proteins. The structures of MHC I and II are discussed in *Chapter 4* in the context of their roles in the immune response. Here their involvement in triggering transplant rejection will be emphasised. Proteins encoded by *MHC I* are single polypeptides (monomers) that are always associated with a smaller protein, β_2 microglobulin (β_2M), encoded on chromosome 15. However, MHC II proteins are, αβ heterodimers, with both polypeptides encoded by *MHC II*. A simplified illustration of the HLA region is shown in *Figure 6.22*.

The HLA complex contains a number of genetic loci, including those that encode different types of class I proteins. Thus, the HLA-A, -B and -C regions contain genes that encode HLA-A, -B and -C proteins, respectively. These are found on all nucleated cells and are not allelic forms but distinct types of proteins. However, there *are* allelic forms of each of the *HLA-A*, *-B* and *-C* genes, and these encode HLA proteins, which have small differences in their primary structures. Hence, each class I gene is **polyallelic**, meaning many different alleles exist, although, of course, each individual only expresses a maximum of two alleles at each locus (chromosomes occur in pairs; *Chapter 15*). Moreover, the alleles are codominant so that each nucleated cell expresses two different alleles of the *HLA-A* gene, as well as two variants of *HLA-B* and *HLA-C*.

In the HLA complex, the class II region is located nearest to the centromere, which contains the DP, DQ and DR loci. Within each of these loci are the genes encoding the α

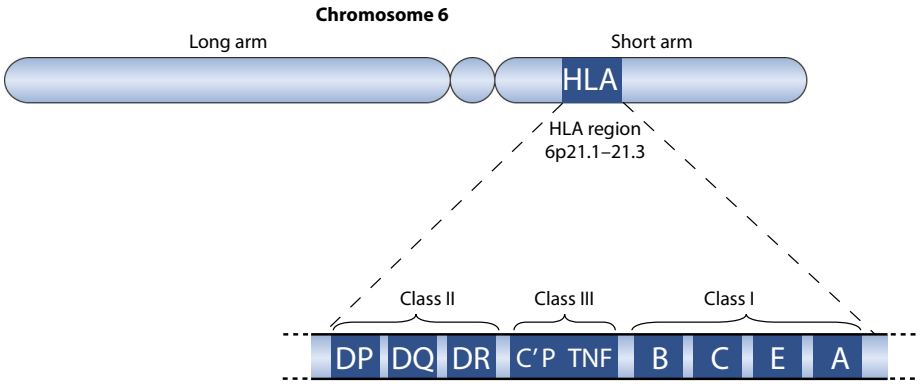


Figure 6.22 Schematic showing a simplified structure of the human leukocyte antigen (HLA) system on chromosome 6.

and β polypeptides of the class II molecules. The arrangements are rather more complex than in the class I region, because each locus may contain more than one gene encoding the α and β chains. For example, the HLA-DR region contains three or four genes for the HLA-β polypeptide. All β gene products may be expressed in a single cell, making the degree of variation much larger than that of class I proteins (*Table 6.13*). Like class I, the class II region also displays a high degree of polymorphism. The HLA complex in humans is the most highly polymorphic system known: nearly 28 400 class I and over 12 500 class II alleles have been identified as of 2025, a huge increase since the first edition of this book was published (*Figure 6.23*)!

Gene	Codes for
HLA-DRB	β chain of HLA-DR
HLA-DQA1	α chain of HLA-DQ
HLA-DQB1	β chain of HLA-DQ
HLA-DPA1	α chain of HLA-DP
HLA-DPB1	β chain of HLA-DP

Table 6.13 Alleles of the class II human leukocyte antigen (HLA) genes.

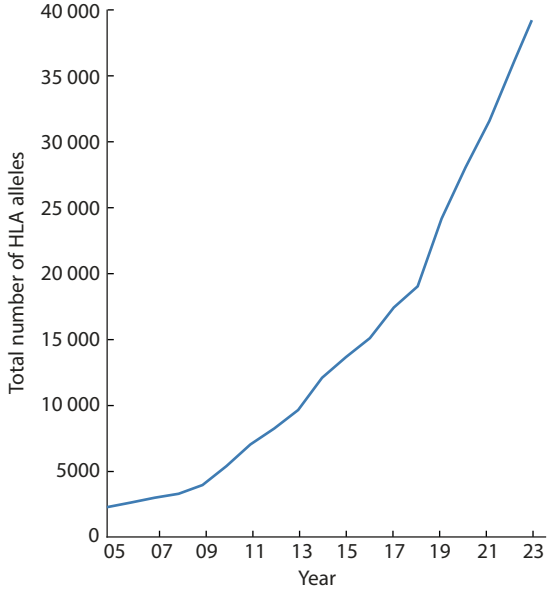


Figure 6.23 An illustration of the increasing number of human leukocyte antigen (HLA) alleles identified since 2005, when the first edition of this book was published. Data from IPD-IMGT/HLA Database.

Margin Note 6.8
Class III proteins

The class III region of the HLA complex is situated between the other two as shown in *Figure 6.22*. It includes genes that encode several secreted soluble proteins found in plasma and other body fluids. These include some of the complement proteins, such as C2, C4 and factor B, and molecules associated with inflammation, for example the cytokines tumour necrosis factor (TNF)-α and -β (*Figure 6.24*) that have immune functions (*Chapter 4*). Class III proteins therefore have different functions from those of class I and II and are not *directly* involved in graft rejection.

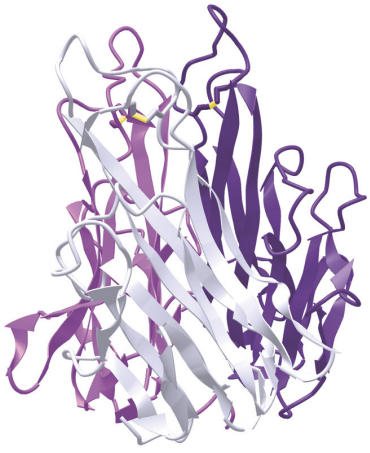


Figure 6.24 Molecular model of a tumour necrosis factor (TNF). PDB file 1TNF.

Given that each individual has two of each of these alleles and that the allelic forms are codominant, it is apparent that the chances of two unrelated individuals having identical HLA genes is vanishingly small, making the possibility of finding a 'good' transplant match, other than between identical twins, correspondingly low.

HLA typing

Human leukocyte antigen typing is the determination of the HLA proteins on the cells of both the recipient and the potential donor. Identifying the HLA types of potential transplant recipients and storing this information on computer databases makes it possible to match donor organs that become available to the most appropriate waiting recipient.

The degree of match required depends on the organ being transplanted and is much more stringent for HSCTs (see *Box 5.2* and *Section 6.14*) than for solid organ transplants. The transplantation of HSCs is used to treat a number of conditions, including immunodeficiency diseases and some kinds of cancer (*Chapters 5* and *17*). Transplantation of bone marrow presents particular difficulties because it contains immunocompetent cells that can lead to a fatal GVHD.

The traditional method for HLA typing is a serological technique that uses antibodies to known HLAs. This method may still used but is being superseded by molecular biological techniques, which give greater accuracy and resolution in determining HLA differences and similarities with respect to the donor and recipient.

Serological HLA typing

The lymphocytotoxicity assay is used in serological HLA typing. The cells to be typed are peripheral blood mononuclear cells (PBMCs), which are readily obtained from whole blood. For typing class II HLAs, it is necessary to use purified B lymphocytes, since resting T lymphocytes do not express class II molecules. The cells to be typed are analysed by automated robotic devices. Aliquots of the cell suspensions are pipetted into the wells of 96-well trays, known as Teresaki plates, named after the inventor of this test (Paul Teresaki (1929–2016); *Figure 6.25*). Antibodies to individual HLAs are added to each well and will bind to the lymphocytes if they express the appropriate antigen. The addition of complement to all the wells results in the deaths of those cells with bound antibodies. Viability stains, such as fluorescent acridine orange and ethidium bromide, are then used to reveal which wells contain dead cells. Acridine orange enters living cells and stains the nuclei green, while ethidium bromide penetrates dead cells staining their nuclei red (*Figure 6.26*).

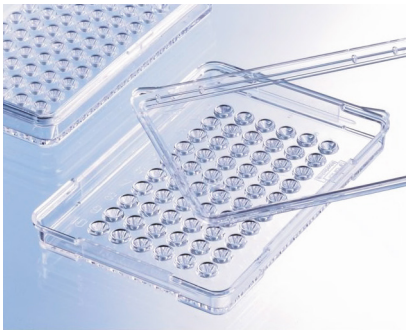


Figure 6.25 A Teresaki plate used in serological human leukocyte antigen (HLA) typing.

The anti-HLA antibodies used in these tests can be obtained from people already sensitised to HLAs or they may be monoclonal antibodies with specificities for known HLAs. Rabbit serum is used as a source of complement. A test is scored as strongly positive when more than 50% of the cells in a well are killed.

A serological cross match detects the presence of antibodies to graft antigens in the serum of a potential recipient. Serum from the potential recipient is added to PBMCs (lymphocytes and monocytes) from the donor. Complement is added and the viability of the cells

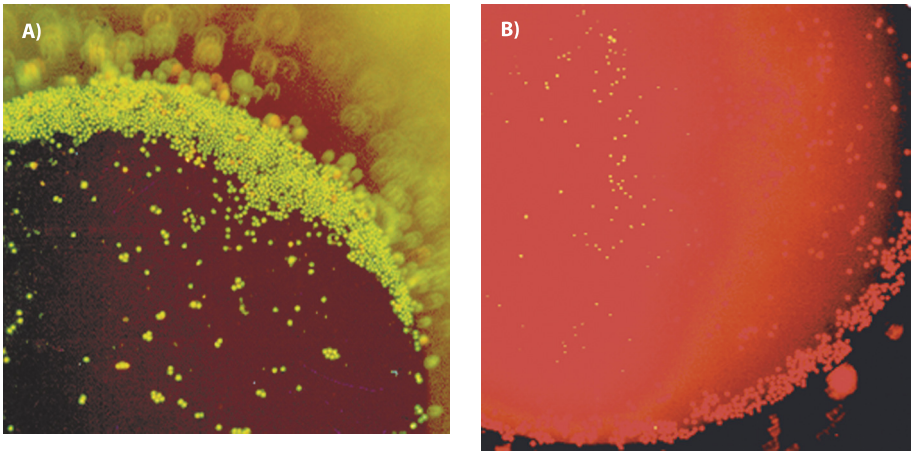


Figure 6.26 Serological test to determine human leukocyte antigen (HLA) type. (A) A negative result, that is, no cell deaths and (B) a positive result in which cells have been killed as indicated by different fluorescent colours. Courtesy of the Transplantation Laboratory, Manchester Royal Infirmary, UK.

tested as previously. If donor cells are killed then it is probable that the recipient already has antibodies against graft antigens and the donor is unlikely to be a suitable match.

Molecular biology HLA typing

The main molecular biological technique used in histocompatibility testing is the PCR (*Chapter 3*). Samples of DNA from the individual being typed are amplified many times over by the PCR; therefore, initially, only a small amount of DNA is required. Living cells are not needed because the PCR can be carried out on whole blood that has been stored frozen. In the sequence-specific primer (SSP) assay, the primers used are those specific for individual HLA alleles. The DNA will only be amplified in those mixtures that contain a probe complementary to that of the DNA being typed (*Figure 6.27*).

In the sequence-specific oligonucleotide probe (SSOP) method, all the DNA is amplified by the PCR. Sequence-specific oligonucleotide probes are then used to identify the products. In some laboratories, the oligonucleotide probes are attached to microspheres. Incubation of the microspheres with the PCR product allows the latter to bind to any complementary HLA sequences. The microspheres, with bound DNA, are then analysed by flow cytometry to identify target HLA sequences.

Matching donor and recipient

Individuals inherit a set of HLA genes from each parent. The greater the similarity between the HLA genes of donor and recipient, the more likely the transplant is of being successful. Thus, for example, transplants between siblings are more likely to have a closer match than between unrelated donors. The HLA profiles of potential recipients are stored in computer databases so that, when an organ becomes available, it can be given to a waiting patient whose HLAs match it as closely as possible. Retrospective studies on the outcome of kidney transplants have shown that matching the donor to the recipient improves graft survival. Thus, in the most favourable match, that is, no mismatches of the *HLA-DR*, *HLA-A* and *HLA-B* alleles, graft survival is superior to those with no mismatch of *HLA-DR* and only one mismatch of *HLA-A* and/or *HLA-B* alleles. These also give significantly better graft survival than any other mismatched graft.

The UK operates a database of individuals who have volunteered to act as donors in the event of their deaths (*Figure 6.28*). However, there is always a shortage of donor organs, and patients often have extremely long, debilitating waits before a suitable one becomes available.

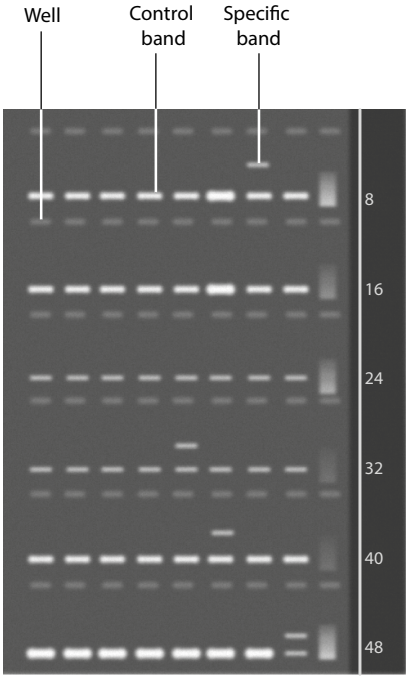


Figure 6.27 Polymerase chain reaction (PCR) assay using sequence-specific primers (SSPs) to determine human leukocyte antigen (HLA) type. The specific band shows that the individual being typed is positive for that specific HLA gene. The controls ensure that the patient's DNA has been amplified correctly. The numbers down the right-hand side are reference values to indicate the appropriate sizes of the bands. Courtesy of the Transplantation Laboratory, Manchester Royal Infirmary, UK.

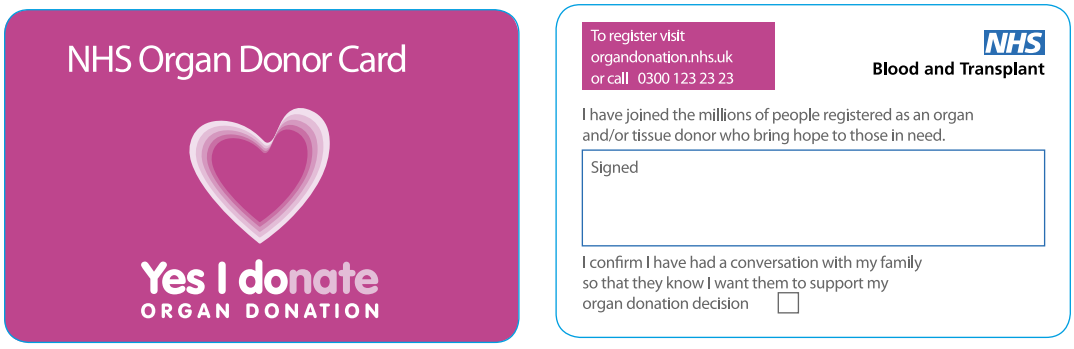

NHS Organ Donor Card
Yes I donate
ORGAN DONATION
To register visit
organdonation.nhs.uk
or call 0300 123 23 23
NHS
Blood and Transplant
I have joined the millions of people registered as an organ and/or tissue donor who bring hope to those in need.
Signed
I confirm I have had a conversation with my family so that they know I want them to support my organ donation decision

Figure 6.28 Front and back views of a UK organ donor card.

All patients who receive an allograft are liable to reject the transplant, even if the recipient and donor are closely matched for histocompatibility antigens. This is because relatively few of the HLAs are tested for and complete matches are rare. Thus, all patients who receive an allograft have to take immunosuppressive drugs to prevent rejection resulting from immune responses.

6.13 IMMUNOSUPPRESSION

The first kidney transplant by Murray in the 1950s, although surgically successful, was, unfortunately for the patient, a medical failure because of transplant rejection. A later kidney transplant between identical twins (an isograft) was, however, clinically successful, which emphasised the role of the immune system in organ rejection. Fortunately, immunosuppression drug regimens have been developed to combat rejection. Indeed, Murray was among the pioneers in using an immunosuppressant

combination therapy of corticosteroids and azathioprine (AZA) to prevent the rejection of transplanted kidneys.

Immunosuppressant drugs used to prevent rejection of the tissue/organ transplant must be administered with care, since they also depress the ability of the patient to resist infectious agents (*Chapters 2* and *3*) and often have adverse side effects.

All the 'first generation' drugs produce a 'blanket' immunosuppression and depress all immune responses. This makes patients more susceptible not only to infections of all kinds, but especially to opportunistic infections by organisms such as *Candida albicans*. Immunosuppressed patients are also more susceptible to cancers associated with viruses. These include lymphomas and Kaposi sarcoma, which are associated with the Epstein–Barr virus (EBV) and Kaposi sarcoma-associated herpes virus (KSHV), respectively (*Chapter 17*). First-generation immunosuppressive treatments also have significant toxicity because they affect all dividing cells, including those of the bone marrow and of the gastrointestinal tract (GIT). Others, such as methotrexate, also show liver toxicity.

Immunosuppression therapy has advanced greatly since these early treatments both in the agents used and in the understanding of their actions at the molecular level. However, a delicate balance in the amount of drug given must still be achieved to prevent rejection but not leave the patient too susceptible to infection and debilitated from drug toxicity. Multidrug combinations are always used because this allows lower amounts of any one agent to be used, and their different, often complementary, antirejection mechanisms increase the chances of success.

Immunosuppressive drugs are used therapeutically for:

- induction
- maintenance
- treatment of rejection

Induction involves a 'full on' medical approach immediately following surgery, when the risks of rejection are deemed highest. Maintenance therapy may involve the same or other combinations of agents but in reduced amounts. Similarly, treatments combating rejection may use some of the same drugs at increased doses or other combinations of immunosuppressants.

Numerous immunosuppressants were first developed to combat immunological disorders (*Chapter 5*) and cancers (*Chapter 17*) before being adapted for use in transplantation; others were found fortuitously. *Table 6.14* lists examples of immunosuppressive drugs and outlines their uses. It is apparent from *Table 6.14* that the clinical applications of immunosuppressants overlap; however, their structures and modes of actions differ greatly and will now be briefly discussed.

Margin Note 6.9
Murray and the Nobel Prize

For performing the first successful human organ transplant and his research on the uses of radiation and drugs, especially azathioprine (AZA), to prevent graft rejection, Murray was jointly awarded the 1990 Nobel Prize in Physiology or Medicine.

Immunosuppressant drug	Clinical uses (usually in combination with other agents)
Corticosteroids, for example prednisolone	Induction, maintenance, rejection therapy
Purine analogues, for example azathioprine	Induction, maintenance (largely replaced by other agents)
Alkylating agents, for example cyclophosphamide	Induction, maintenance (largely replaced by other agents)
Enzyme inhibitors, for example mycopophenolate mofetil	Maintenance, rejection therapy
Folic acid (folate) antagonists, for example methotexate	Mainly used in cancer therapy
Calcineurin inhibitors, for example cyclosporin, tacrolimus	Induction, maintenance, rejection therapy
Cell cycle antagonists, for example sirolimus	Induction, maintenance
Antibodies, for example antithymocyte globulins, monoclonal antibodies	Induction, rejection therapy

Table 6.14 Examples and uses of some immunosuppression drugs. ⏎

Corticosteroids, such as prednisolone (*Figure 6.29*), hydrocortisone and dexamethasone, are immunosuppressive agents that act principally to lower the concentrations of circulating cytokines and reduce inflammation. Despite being among the first generation, they are still extensively used, usually in combination with other drugs such as tacrolimus (see below).

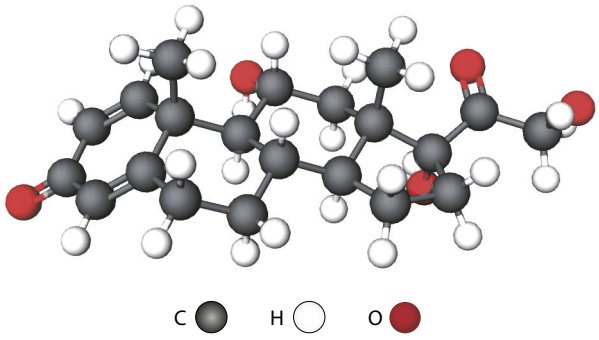


Figure 6.29 Computer drawn model of prednisolone.

Purine analogues, such as AZA and mercaptopurine (*Figure 6.30*), inhibit DNA biosynthesis. Azathioprine is a prodrug that is converted to mercaptopurine *in vivo*. Mercaptopurine is incorporated into the newly synthesised strands of DNA during replication and stops their elongation, and so reduces the proliferation of T cells. Alkylating agents, such as cyclophosphamide, also impair DNA biosynthesis but by reacting with its strands and cross-linking them together, so preventing further replication.

Like AZA, mycophenolate mofetil is a prodrug that is metabolically converted to mycophenolate (MPA; *Figure 6.31*), a potent inhibitor of inosine monophosphate dehydrogenase (IMPDH). Inosine monophosphate can be thought of as having a central role in the *de novo* biosynthesis of the purine nucleotides dATP and dGTP that are essential for nucleic acid formation. Most cells also possess a salvage pathway able to synthesise purine nucleotides from the free purine bases obtained from digested foods or by the catabolism of nucleic acids (*Figure 6.31*). Inhibition of IMPDH by MPA restricts the biosynthesis of guanine nucleotides and reduces the proliferation of both B and T lymphocytes, which lack the salvage pathway present in other cells. Methotrexate also decreases the biosynthesis of nucleic acids but by inhibiting dihydrofolate reductase, an enzyme involved in the biosynthesis of purine and pyrimidine bases. The drug resembles the vitamin folic acid/folate (*Chapter 10*) in structure (*Figure 6.32*), which forms an essential cofactor for the enzyme. The antagonistic action of methotrexate against folic acid inhibits the biosynthesis of purine and pyrimidine bases and therefore that of nucleic acids.

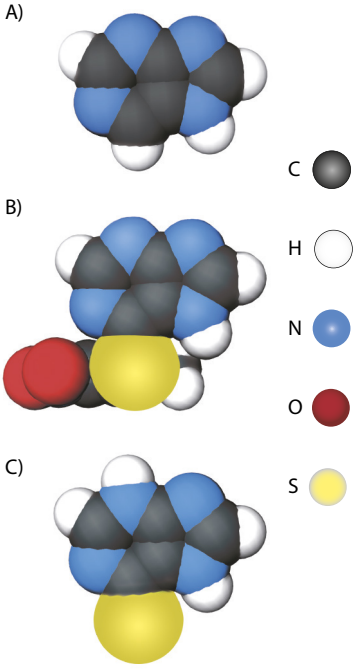


Figure 6.30 Computer drawn models of (A) purine, (B) azathioprine and (C) mercaptopurine. Note the similarities in the structures of the three compounds.

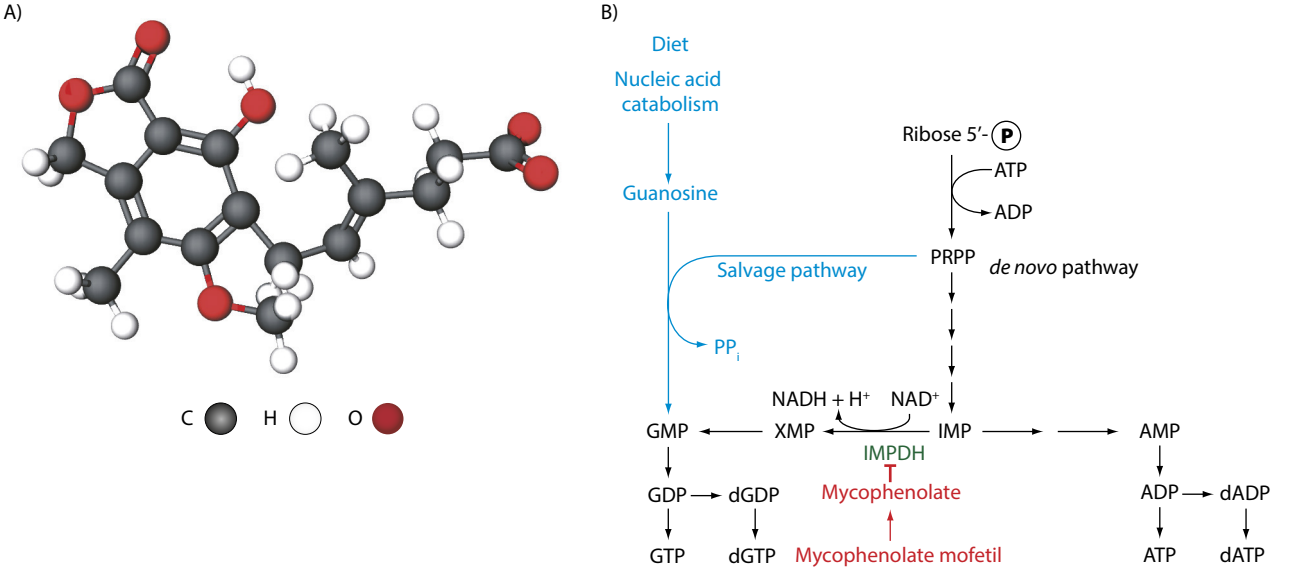


Figure 6.31 (A) Computer drawn model of mycophenolate. (B) Outline of the biosynthesis of purine nucleotides and its inhibition by mycophenolate.

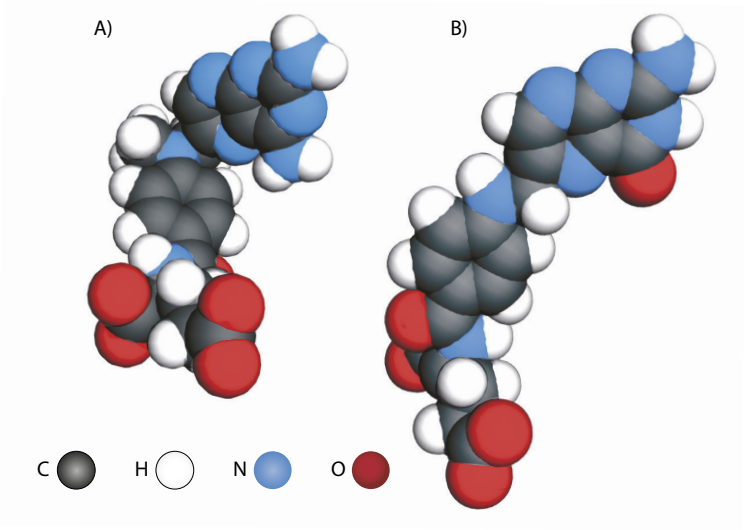


Figure 6.32 Computer drawn models of (A) methotrexate and (B) folate. Note the resemblance in their structures.

Cyclosporin and tacrolimus are major clinical immunosuppressants. They function by inhibiting the catalytic activity of the cytosolic phosphatase calcineurin (CaN; *Figure 6.33*). Calcineurin plays an essential role in several signalling pathways that use Ca^{2+} as a second messenger (*Chapter 7*). Binding of the first messenger, for example a cytokine or hormone, to the extracellular portion of its receptor increases the cytosolic concentration of Ca^{2+}, which, in turn, binds to the Ca^{2+}-sensing protein calmodulin (CAM; *Figure 6.34*). The Ca^{2+}–CAM then forms a Ca^{2+}–CAM–CaN complex, which is the active phosphatase. Active CaN has over 200 substrates but the most relevant here are a family of *nuclear factors of activated T cells* (NFATs). In their phosphorylated form in the cytosol, NFATs are quiescent; however, their catalysed dephosphorylation by the Ca^{2+}–CAM–CaN complex exposes a previously hidden nuclear localisation signal that directs the translocation of the factor through nuclear pores into the nucleus (*Chapter 16*). Here, it binds to specific genes, altering their expression and upregulating the production of several cytokines including interleukin-2 (IL-2), leading to T cell activation.

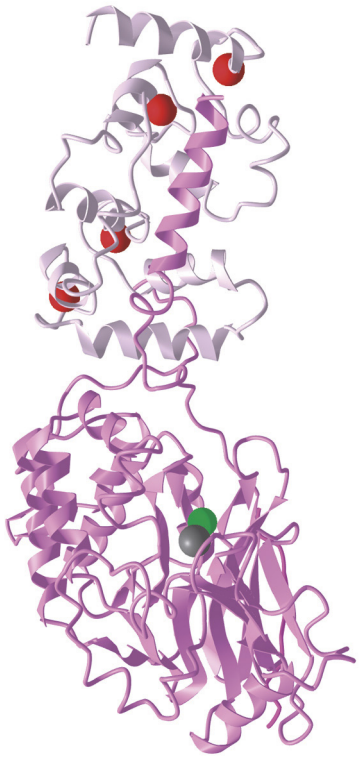


Figure 6.33 Molecular model of calcineurin. The Zn–Fe centre at the active site is shown as grey- and green-coloured spheres, respectively. The four bound Ca^{2+} are shown in red. PDB file 3LL8.

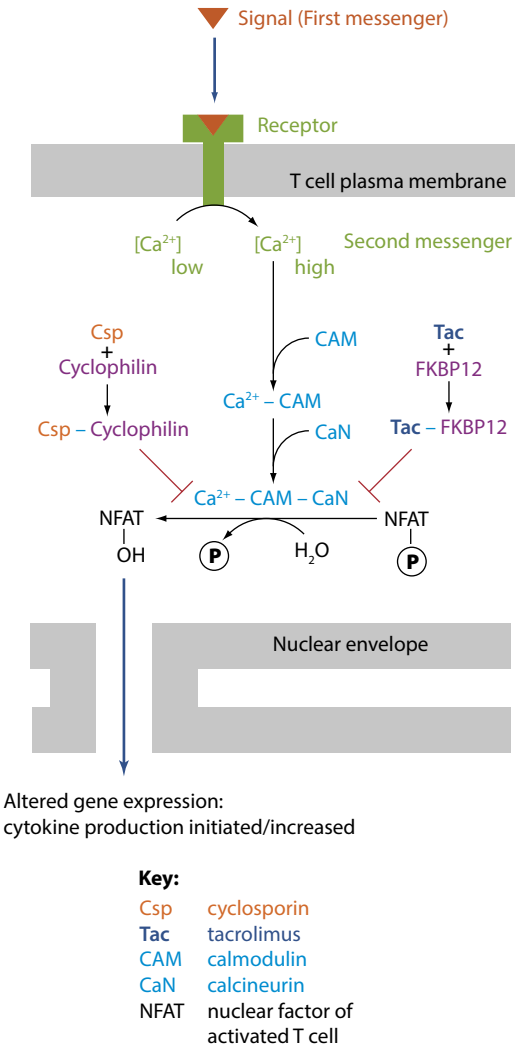


Figure 6.34 Outline of the activation and role of calcineurin and its inhibition by cyclosporin and tacrolimus. Movements of molecules are shown as blue arrows.

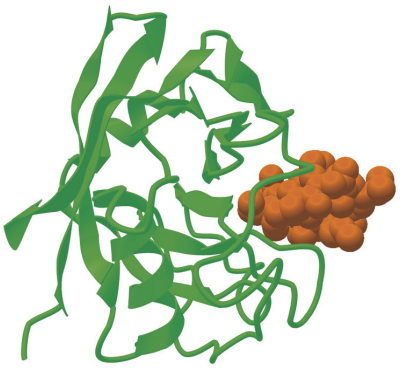


Figure 6.35 Molecular model of a cyclosporin–cyclophilin (coloured red and green, respectively) complex. PDB file 1CWA.

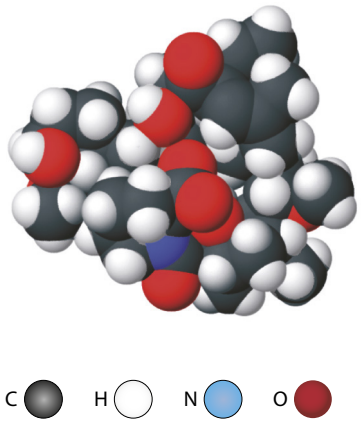


Figure 6.36 Computer drawn model of tacrolimus.

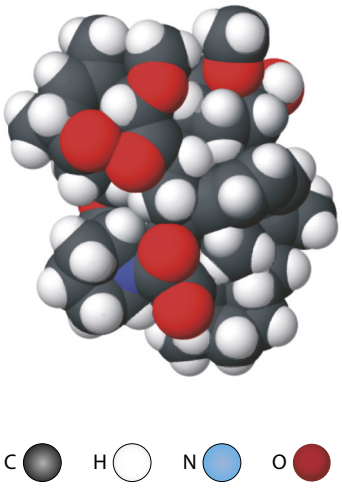


Figure 6.37 Computer drawn model of sirolimus.

Cyclosporin and tacrolimus are major clinical immunosuppressive drugs because they inhibit calcineurin, although by different molecular mechanisms (*Figure 6.34*), and so impair T cell activation. Cyclosporin A is a cyclic peptide derived from the soil fungus *Tolypocladium inflatum gams*. It has considerable immunosuppressive action, without the bone marrow toxicity demonstrated by some other drugs. Cyclosporin binds to the cytoplasmic immunophilin protein cyclophilin, and it is this complex (*Figures 6.34* and *6.35*) that inhibits the phosphatase actions of the Ca^{2+}–CAM–CaN complex. Tacrolimus, a macrolide antibiotic (*Figure 6.36*) derived from *Streptomyces tsukubaensis*, resembles cyclosporin A in its inhibitory action but binds to the immunophilin FK506-binding protein 12 (FKBP12; *Figure 6.34*). It is the tacrolimus–FKBP12 complex that inhibits CaN activity. This is a more potent inhibitor than the cyclosporin complex but also has more side effects.

Sirolimus (*Figure 6.37*), also known as rapamycin, is produced by the actinomycetes *Streptomyces hygroscopicus*. It also binds to FKBP12, but the complex does not inhibit CaN activity. The metabolic target of the sirolimus–FKBP12 complex is, rather bizarrely, called the *mechanistic target of rapamycin* (mTOR), which was previously known as the mammalian target of rapamycin. This target is a highly conserved protein kinase found in two distinct protein complexes. One, the TOR complex-1 (TORC1), is a key regulator of cell growth and proliferation, and mRNA translation. The second, TORC2, promotes actin cytoskeletal rearrangements, cell survival and cell cycle progression (*Chapters 15* and *16*). The sirolimus–FKBP12 complex exerts its immunosuppression effects by binding to TORC1 and inhibiting cell cycle progression at the G1-S transition stage and by impairing the IL-2-dependent progression of T cells through the cycle.

Antilymphocyte globulins (ATG) are a mixture of antibodies raised against human T lymphocytes, which causes their lysis by both complement and cell-mediated opsonisation (*Chapter 4*). It was developed in the 1970s by injecting human lymphoid cells into rabbits or horses and then collecting and purifying the resulting polyclonal antibodies. Rabbit preparations are generally favoured as they cause fewer side effects when used for induction and in treating acute solid organ rejection. Its use has been restricted since the development of monoclonal antibodies (MAbs) in the 1980s. Those MAbs used clinically are either human or humanised mouse types (*Chapter 4*). The latter are chimeric molecules produced by replacing the mouse variable region with that of a human antibody. The advantages of using humanised MAbs is that specificity is retained without the antigenicity of the original mouse molecule and so they can be used in long-term treatment regimens. The clinical use of MAbs means that the molecular activities of B and T lymphocytes and other immune cells, such as APCs, and the interactions between immune cells can be specifically targeted. Many MAbs have been developed for clinical use to interfere with immune functions and prevent or treat organ rejection. Here, three will be described to illustrate their range and activities: anti-CD3, anti-CD2 and anti-CD25.

The protein CD3 is a hexamer of four different types of transmembrane polypeptides that form a complex with the T cell receptor (TCR; see *Figures 4.34* and *6.38*). When a T cell docks with an APC carrying an antigenic peptide, the initial interaction is stabilised by the CD4 molecule of the T cell interacting with the MHC class II protein of the APC. The cell adhesion molecule CD2 is also an integral membrane protein found on the surfaces of T cells and natural killer cells (NKCs). It interacts with other adhesion molecules, such as lymphocyte function-associated antigen-3 (LFA3), which is expressed on the surfaces of other cells, such as APCs. The associations of the TCR-CD3-CD4 and other adhesion molecules, such as CD2 and LFA3, have been referred to as an immune synapse. The formation of this molecular complex connects the two cells and initiates an extended series of signalling events within the T cell that result in its complete activation.

Both anti-CD3 and anti-CD2 MAbs are effective immunosuppressants: by blocking T cell functions and proliferation, they have proved to be effective medications in treating acute rejections. A mouse anti-CD3 was the first MAb to be approved for clinical use in 1986 by the US FDA and by European authorities shortly after. Unfortunately, it became apparent that its use was associated with cytokine release syndrome (CRS, 'cytokine storms'; see *Chapter 4*) and its manufacturers withdrew it in 2010.

The α subunit of the IL-2 receptor or CD25 (see *Figure 4.4*) is only expressed on non-resting and activated T cells. The binding of an anti-CD25 MAb to the receptor is therefore immunosuppressive because it inhibits the proliferation of T cells.

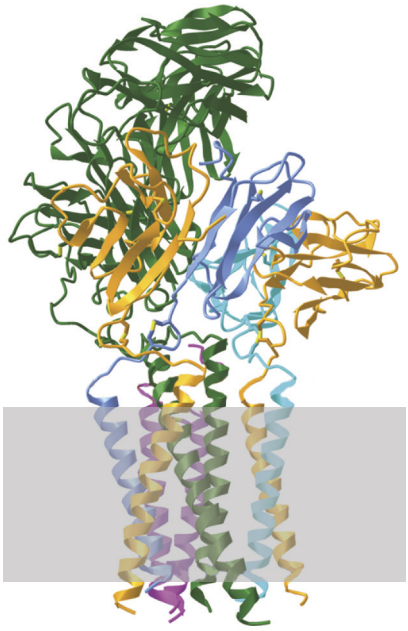


Figure 6.38 Molecular model of the CD3–TCR complex. The two subunits of the receptor are shown in dark green, while the CD3 ζ monomers are magenta, its ε monomers are orange and the δ and γ monomers are cyan and light blue, respectively. The grey block indicates the approximate position of the plasma membrane of the T cell. PDB file 6JXR.

6.14 HAEMATOPOIETIC STEM CELL TRANSPLANTATION

Transplantation of HSCs is used to correct some immunodeficiency diseases (*Chapter 5*) and to treat certain types of cancers, such as leukaemias and lymphomas (*Chapter 17*). Patients with total bone marrow failure, such as in aplastic anaemia, or partial failure, as seen in some of the haemoglobinopathies (*Chapter 13*), may also benefit from an HSCT.

Haematopoietic stem cells for transplantation may be derived from peripheral blood or bone marrow and from umbilical cord blood (see *Box 5.2*). Transplantations may be:

- **autologous**
- **syngeneic**
- **allogeneic**

In an autologous transplant, the patient's stem cells are harvested prior to treatments in which the patient's own stem cells are subsequently destroyed, for example by radiation or high-dose chemotherapy to treat leukaemia. The harvested stem cells are stored in liquid nitrogen (–196°C) until the transplant can take place, when they are carefully thawed and reinfused into the patient.

A syngeneic transplant is an HSCT between identical twins. Given their identical genotypes, this type of transplant involves little chance of rejection. In contrast, an allogeneic HSCT involves donors and recipients who are genetically nonidentical. The donor could be a sibling or a potential donor identified in a tissue registry (*Margin Note 6.10*). However, donors who are related to the recipients are more likely to be compatible.

There is a greater need for close HLA matching of donors and recipients in HSCTs, as GVHD occurring 7–30 days post-transplantation is a major hazard. Acute GVHD implies a clinical failure of immunosuppression therapy (see *Section 6.13*). Epithelial cells in the skin and those lining the intestine are attacked by sensitised T cells originating in the graft. Patients present with a severe maculopapular skin rash

Margin Note 6.10
HSC databases

Potential HSC donors are always required to increase the number of HLA types available. In the UK, the main registry, the British Bone Marrow Registry (BBMR), works cooperatively with the other UK bone marrow and blood donor registries, namely the Anthony Nolan charity and the NHS Cord Blood Bank. Donors need only register with one UK agency, as they share matching information. Similarly, in the USA, the National Marrow Donor Program (formerly Be the Match) maintains an international registry of stem cell donors.

(*Margin Note 6.11*) and sloughing of the intestinal epithelium that may lead to diarrhoea. Why the skin and intestine are initially subject to greater damage than other tissues and organs is not fully understood (*Figure 6.39*). Splenomegaly and hepatomegaly occur, as these organs become attacked by T lymphocytes and the patient may become jaundiced. Acute GVHD is frequently fatal. A chronic form with similar symptoms may occur over a longer time frame. The patients may suffer frequent secondary infections. Once GVHD develops, it must be treated with increased doses of immunosuppressive drugs and/or alternative ones to those used in induction and maintenance.

One method of avoiding GVHD is to remove T lymphocytes from HSCTs prior to reinfusion. This process, known as T cell depletion, involves using antibodies against the T cells.

Margin Note 6.11 ⏎ Maculopapular skin rashes

Macules are flat reddened areas of skin that present as a rash, while papules are raised areas of skin in a rash. Thus, a maculopapular rash is one that presents with both flat and raised parts.

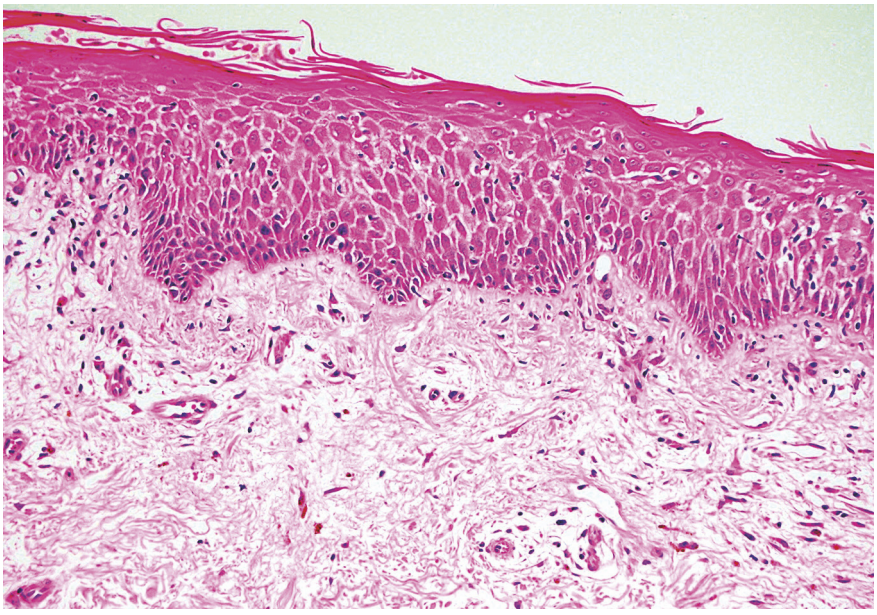


Figure 6.39 A biopsy photomicrograph showing the characteristic skin damage caused by graft versus host disease (GVHD). The epidermis is spongiotic, that is, the keratinocytes are separated by intercellular oedema, and contains numerous donor-derived T lymphocytes (identifiable by their dark nuclei) between the keratinocytes. Some of the T cells are adjacent to an apoptotic keratinocyte, a histological feature called satellite cell necrosis that is characteristic of cutaneous GVHD. Photograph courtesy of Dr Jon Salisbury, Consultant Histopathologist, King's College Hospital, UK. ⏎

Peripheral blood stem cell transplantation

Peripheral blood stem cell transplantation (PBSCT) is the most common form of HSCT. The collection of stem cells from peripheral blood is easier for both the collector and the donor. In addition, engraftment of a PBSCT is often more rapid than with bone marrow. The donor is treated with granulocyte colony-stimulating factor (G-CSF; see *Figure 5.5* in *Box 5.2*) to increase the number of stem cells in the blood. The stem cells are then obtained from the donor by leukapheresis (see *Margin Note 6.6*). Leukapheresis may take several hours to complete, and more than one session may be needed in 10% of all PBSC donations, with stem cells being stored frozen between donations.

Bone marrow transplantation

The first successful bone marrow transplantation (BMT) was achieved in 1968. Bone marrow contains the HSCs that give rise to all the formed elements of the blood (see *Figure 13.2*). In BMTs, the donor receives a general or local anaesthetic and the marrow is harvested using a needle inserted through the skin over the pelvic bone and into the bone cavity. The process takes approximately one hour to harvest sufficient quantity; usually a minimum of 3×10^8 nucleated bone marrow cells containing approximately 2×10^6 stem cells per kilogram of recipient body weight is required. The marrow is treated to remove blood and plasma, especially if there is a disparity between the ABO blood groups of the donor and the recipient. Fragments of bone are removed and the marrow

may also be depleted of T cells to reduce the risk of GVHD prior to intravenous transfusion into the recipient. The recipient may be given antibiotics because, until the marrow is engrafted and starts to produce blood cells, the patient is at risk of infection. They may also be given platelet and erythrocyte transfusions to prevent bleeding and anaemia. A patient receiving a BMT may show adverse side effects, such as nausea, fatigue, hair loss and reduced appetite.

Umbilical cord stem cell transplantation

Stem cells may also be obtained from umbilical cord blood, of course with the permission of the family involved. Following birth, blood derived from the baby is obtained from the umbilical cord and the placenta. Since only a small amount of blood is retrieved in this way, the collected stem cells are typically used to treat children.

Identification of HSCs

Haematopoietic stem cells can be identified by the presence of CD34, a transmembrane glycoprotein. Thus, the numbers of CD34+ cells, that is, HSCs, in a sample can be assessed by staining it with a fluorescently labelled antibody to CD34. Cell numbers can then be determined by fluorescence microscopy or flow cytometry (see *Box 6.1* and *Figure 6.40*). Both peripheral and cord blood may be further processed to obtain the stem cells. For example, an anti-CD34 antibody linked to magnetic particles will bind to CD34+ cells, which can then be purified using a magnet.

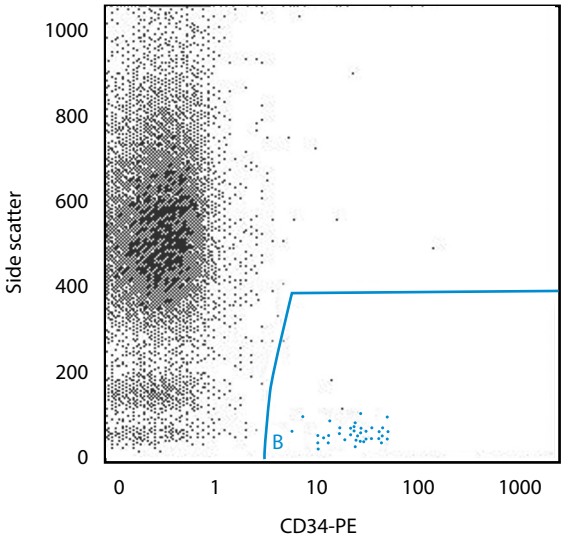


Figure 6.40 Flow cytometric characterisation of CD34+ stem cells (isolation in the gated area B) from umbilical cord blood. Courtesy of Dr T.F. Carr, Royal Manchester Children's Hospital, UK.

Storage of transplant material

Harvested stem cells may be stored for two to three days in a refrigerator at 4°C. This may be required, for example, if more than one harvesting procedure is needed or if a patient has to undergo radiation or chemotherapy for cancer treatment (*Chapter 17*) prior to receiving the graft. If longer term storage is required, the stem cells are stored in liquid nitrogen. Prior to storage, the cryopreservative dimethylsulphoxide (DMSO; *Figure 6.41*) is added to prevent ice crystal formation, which would destroy the cells. A programmed freezer allows the cells to be cooled at the optimum rate for cell survival, which is normally approximately –1°C/min.

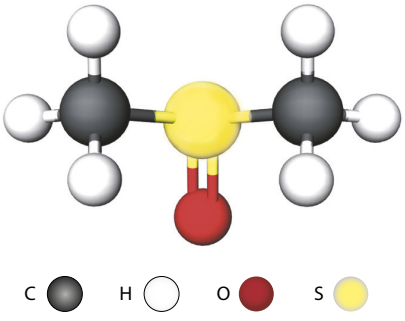


Figure 6.41 Dimethylsulphoxide.

Post-transplant support

Patients who have just undergone PBSCT or BMT are initially highly susceptible to infections. If transfusions of blood products are required, for example for anaemia, these products should be free of CMV (see *Figure 6.18*), a virus which is latent in some individuals. In an immunosuppressed individual, CMV can cause serious illness.

6.15 CASE STUDIES

CASE STUDY 6.1

Marie, a 31-year-old florist, is blood group O, RhD negative, and seven months' pregnant. The father of her baby is blood group A, RhD positive. Marie is worried about the pregnancy, as she has had several miscarriages, each of which occurred before three months of gestation. Despite these miscarriages, Marie has never become sensitised to the D antigen, and anti-D antibodies are *not* present in her circulation.

Questions

(a) What are the chances of her baby developing HDN?
(b) Is Marie likely to become sensitised to the D antigen after the birth of this child?

CASE STUDY 6.2

John, a 50-year-old school teacher, has recently been treated for acute myeloid leukaemia. His consultant recommended he receive aggressive chemotherapy followed by a stem cell transplant. John is likely to die without this treatment. He has no living siblings who could donate HSCs and both his parents are dead.

Questions

(a) What would be the best approach to treating John?
(b) Is an HSCT feasible?

CASE STUDY 6.3

Michael was taken to the accident and emergency unit of his local hospital following a stabbing incident. Michael had lost over a litre of blood and required an immediate transfusion. He was given compatible leukodepleted blood and fresh frozen plasma. However, he suffered acute respiratory distress approximately four hours after the transfusion. He had hypotension, cyanosis and a temperature of 39.5°C. Examination of the chest indicated fluid in his lungs.

Question

(a) What is the likely cause of Michael's respiratory problems?

CASE STUDY 6.4

Jason, an 18-year-old with excellent GCSE and A-level results, recently began a career in publishing. Previously, he had suffered only the usual run of childhood illnesses but, recently, had developed an unhealthy-looking pallor and pronounced red patches on his skin (petechiae), suffered frequent nosebleeds and headaches, and was always too tired to play football for his favourite team. On examination, his general practitioner found the inner lining of the lower eyelid and oral cavity extremely pale but no other obvious physical problems. Blood tests showed:

Result		Reference values
Haemoglobin:	90	130–180 g/L
Erythrocytes:	2.5×10^{12}	$4.5–6.5\times10^{12}$/L
Leukocytes:	1.9×10^{9}	$4.0–11.0\times10^{9}$/L
Platelets:	15×10^{9}	$150–400\times10^{9}$/L

A subsequent bone marrow biopsy showed hypocellular marrow (6% reference numbers) consisting largely of adipose material. This, together with other tests, gives a diagnosis of the potentially fatal condition idiopathic severe aplastic anaemia. Jason was treated with platelet and packed erythrocyte transfusions and appropriate prophylactic antibiotics; however, the condition can only be cured with an HSCT from a suitable donor. Unfortunately, Jason had no siblings. However, a maternal-side cousin was a suitable match. The transplant appeared successful, and Jason was discharged from hospital after 22 days. At 28 days he was readmitted, discharging copious volumes of watery bloody diarrhoea and with extensive skin rashes that had begun on his face but spread rapidly. He was initially treated with corticosteroids and tacrolimus, but the rash persisted and the diarrhoea became worse. The treatment was changed to mycophenolate mofetil and intravenous injections of an anti-CD2 MAb. This regime improved and then eliminated his symptoms. After several months in hospital, Jason was allowed home.

Questions

(a) What condition is characterised by extensive skin rashes and watery bloody diarrhoea?
(b) What class of histocompatibility alleles are most likely to have led to Jason's post-transplant problems?
(c) Suggest why a sibling would have been a more suitable donor than the cousin.
(d) What was the purpose of first administering corticosteroids and tacrolimus?
(e) What component of the immune system is neutralised by the anti-CD2 MAb?
(f) Why was it necessary to inject the anti-CD2 MAb?
(g) Would you recommend any further medication in the period following Jason's discharge from hospital?

6.16 SUMMARY

Transfusions of blood and blood products are routine and safe clinical procedures, which rarely cause harm. Transfusion has been greatly facilitated by knowledge of the range of blood group antigens and the conditions under which antibodies to blood group antigens can cause problems. The screening of donors, for example for HIV or hepatitis viruses, has also increased the safety of the procedure. The transfusion of incompatible blood can cause the death of a recipient, but this is a rare occurrence. Transfusion laboratories are involved in ensuring that compatible blood is given to a patient and that all aspects of the procedure are safe. In addition, laboratories monitor the presence of antibodies in pregnant women that might cause problems for a developing fetus.

Transplantation of solid tissues/organs and HSCTs has progressed considerably since the first successful kidney transplant, and a wide range of tissues and organs are now transplanted routinely. The rejection of transplants is minimised by carefully tissue matching donor with recipient and by the administration of immunosuppressive drugs. The use of this therapy must be carefully balanced and monitored to minimise the opposing risks of rejection and infection.

QUESTIONS

1. Which combination(s) of plasma and erythrocytes would result in agglutination?
- **(a)** Group A erythrocytes and group B plasma.
- **(b)** Group O erythrocytes and group O plasma.
- **(c)** Group B erythrocytes and group O plasma.
- **(d)** Group B erythrocytes and group B plasma.
- **(e)** Group AB erythrocytes and group A plasma.

2. Which of the following statement(s) is (are) **TRUE**?
- **(a)** MHC II proteins are found on all nucleated cells.
- **(b)** MHC I proteins are found on APCs.
- **(c)** Complement is a protein complex that lyses bacteria.
- **(d)** The classical pathway for complement activation is initiated by IgA.
- **(e)** An ABO incompatible transfusion may be fatal.

3. Which condition(s) would mean that a blood donor should be removed from the register of donors?
- **(a)** Previous infection with hepatitis virus.
- **(b)** Current infection with HIV.
- **(c)** Pregnancy.
- **(d)** Donor aged 78 years.
- **(e)** Intravenous drug abuser.

4. List the possible causes of a potential transplant patient having anti-HLA antibodies in their plasma.

5. Give ONE reason why PBSCT is preferable to BMT.

6. Which of the following is the odd ONE out?
- **(a)** Tacrolimus.
- **(b)** Cyclosporin.
- **(c)** Mycophenolate.
- **(d)** Prednisolone.
- **(e)** Sirolimus.

7. Which of the following statements are TRUE and which are FALSE?
- **(a)** Landsteiner was awarded his Nobel Prize for work involving the ABO (ISBT 001) and Rh (ISBT 005) blood groups.
- **(b)** People of blood group AO have only a single copy of *A* on chromosome 9.
- **(c)** The products of the ABO blood group genes are glycosyltransferases.
- **(d)** The Rh blood group system consists of six haploid genotypes.
- **(e)** The Kidd blood group antigens are receptors for CXC and CC chemokines.
- **(f)** The Duffy blood group comprises eight antigens and also forms the receptor for malarial parasites.
- **(g)** FUT1 and *FUT2* both catalyse the formation of α,1→2 glycosidic bonds.
- **(h)** FUT3 is α,1→4 fucosyltransferase.
- **(i)** Blood groups can be determined using minigel systems and by PCR methods.
- **(j)** Erythrocytes coated with antibodies are susceptible to attack by the complement system.

8. Which of the following statements are TRUE and which are FALSE?
- **(a)** The first successful kidney transplant was performed by Barnard in 1954, and the first human heart transplant in 1967 by Murray.
- **(b)** Xenogeneic transplants are the most likely type to suffer rejection because of the large genetic disparity between the donor and the recipient.
- **(c)** Class I MHC proteins are present on the plasma membranes of all nucleated cells, and are involved in the recognition of virus-infected cells by the precursors of T_C cells.
- **(d)** An acute rejection of transplanted material is normally initiated by the cell-mediated immunity system.
- **(e)** Cyclosporine and tacrolimus exert their immunosuppressive actions by differential inhibition of the same signalling pathway.
- **(f)** Calcineurin is a phosphorylase.
- **(g)** Mycophenolate mofetil is an active immunosuppressive drug.
- **(h)** Both polyclonal and monoclonal antibodies can be used in immunosuppression.
- **(i)** A syngeneic transplant is an HSCT between identical twins.
- **(j)** In acute GVHD, epithelial cells in the skin and those lining the intestine are attacked by sensitised T cells from the patient, resulting in a severe maculopapular skin rash and sloughing of the intestinal epithelium.

FURTHER READING

Avent ND and Reid ME (2000) The Rh blood group system: a review. *Blood* **95**:375–387.

Avery P *et al*. (2020) Whole blood transfusion versus component therapy in adult trauma patients with acute major haemorrhage. *Emerg Med J* **0**:1–9.

Azzi JR *et al*. (2013) Calcineurin inhibitors: 40 years later, can't live without … *J Immunol* **191**:5785–5791.

Barbarino JM *et al*. (2013) PharmGKB: cyclosporine and tacrolimus pathways. *Pharmacogenet Genomics* **23**:563–585.

Bentata Y (2019) Mycophenolates: the latest modern and potent immunosuppressive drugs in adult kidney transplantation; what we should know about them? *Artif Organs* **44**:561–576.

Chinen J and Buckley RH (2010) Transplantation immunology: solid organ and bone marrow. *J Allergy Clin Immunol* **125** (Supplement 2):S324–S335.

Claeys E and Vermeire K (2019) Immunosuppressive drugs in organ transplantation to prevent allograft rejection: mode of action and side effects. *J Immunological Sci* **3**:14–21.

Combs MR (2009) Lewis blood group system review. *Immunohematology* **25**:112–118.

Creamer TP (2020) Calcineurin. *J Cell Commun Signal* **18**:137–148.

Griffith BP *et al*. (2022) Genetically modified porcine-to-human cardiac xenotransplantation. *N Engl J Med* **387**:35–44.

Issa F *et al*. (2010) Role of T cells in graft rejection and transplantation tolerance. *Expert Rev Clin Immunol* **6**:155–169.

Lei T *et al*. (2022) Genetic engineering of pigs for xenotransplantation to overcome immune rejection and physiological incompatibilities: the first clinical steps. *Front Immunol* **13**:1031185.

Makroo RN *et al*. (2019) Retrospective analysis of forward and reverse ABO typing discrepancies among patients and blood donors in a tertiary care hospital. *Transfus Med*, **29**:103–109.

Morgan BP *et al*. (2017) Molecular cell biology of complement membrane attack. *Semin Cell Dev Biol* **72**:124–132.

Mujahid A and Dickert FL (2015) Blood group typing: from classical strategies to the application of synthetic antibodies generated by molecular imprinting. *Sensors (Basel)* **16**:51.

Radwan J *et al*. (2020) Advances in the evolutionary understanding of MHC polymorphism. *Trends Genet* **36**:298–311.

Rathee P *et al*. (2013) Immunosuppressants: a review. *Pharma Innov* **1**:90–191.

Schwarz HP and Dorner F (2003) Karl Landsteiner and his major contributions to haematology. *Br J Haematol* **121**:556–565.

Springer SA and Gagneux P (2016) Glycomics: revealing the dynamic ecology and evolution of sugar molecules. *J Proteomics* **135**:90–100.

Vlaar AP and Juffermans NP (2013) Transfusion-related acute lung injury: a clinical review. *Lancet* **382**:984–994.

Westhoff CM (2019) Blood group genotyping. *Blood* **133**:1814–1820.

Wiseman A (2016) Immunosuppressive medications. *Renal Immunol* **11**:332–343.

Yates C and Wang Q (2022) Immunological techniques. In: Ahmed N *et al*. (eds), *Biomedical Science Practice*, 3rd edn. Oxford University Press.

Zhang W *et al*. (2019) Recent topics on the mechanisms of immunosuppressive therapy-related neurotoxicities. *Int J Mol Sci* **20**:3210–3232.

Useful websites

https://www.ebi.ac.uk/ipd/imgt/hla
http://www.anthonynolan.org
https://www.transfusionguidelines.org
https://www.isbtweb.org

Disorders of the endocrine system

7

OBJECTIVES

After studying this chapter you should be able to:

- list the different types of hormones, giving examples of each type
- outline the structure and roles of the endocrine system
- describe the general mechanisms underlying hormone actions
- explain how disorders of the endocrine system occur
- discuss the causes, clinical investigations and management of some endocrine disorders

7.1 INTRODUCTION

The endocrine system is one of two major control systems in the body, the other being the nervous system; both help control the activities of the body. It consists of a number of ductless glands (*Figure 7.1*) that produce hormones. Hormones are molecules that circulate in the blood and excite or inhibit the metabolic activity of target tissues or organs. These functions maintain and regulate body functions, such as growth and development, responses to stress and injury, reproduction, homeostasis and energy metabolism (*Figure 7.2*).

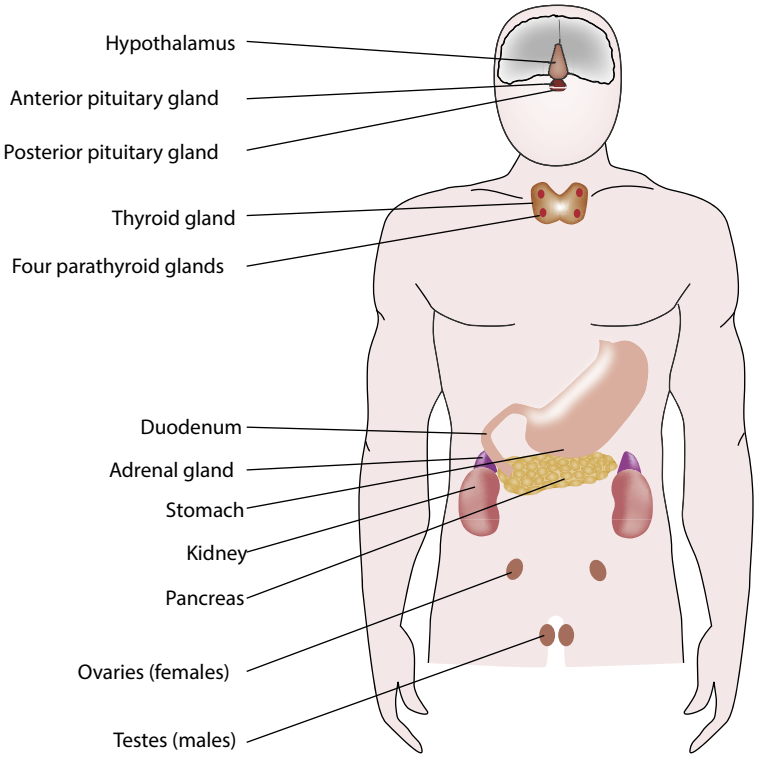


Figure 7.1 The major endocrine glands of the body.

DOI: 10.1201/9781003296836-7

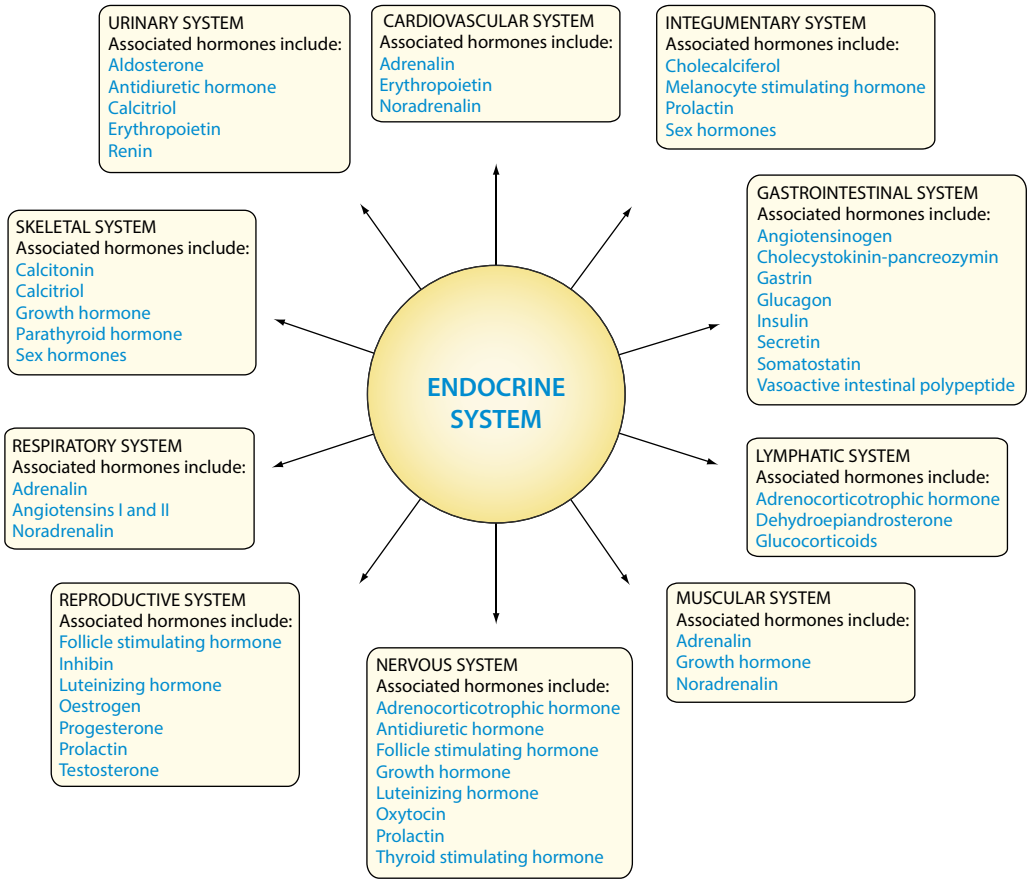


Figure 7.2 Some of the effects of endocrine glands on other body systems. ⏎

Hormones can be divided into three chemical groups: amines, peptides and proteins, and steroids (*Table 7.1*). Many amine hormones, such as adrenaline (epinephrine) and those produced by the thyroid gland, are derivatives of tyrosine. Most hormones are peptides and proteins, examples being insulin and growth hormone. A number of protein hormones, for example thyroid-stimulating hormone, are glycoproteins in that they have glycans (carbohydrate groups) covalently attached to them. Steroid hormones are all derivatives of cholesterol and include cortisol and testosterone. *Figure 7.3* shows examples of each type of hormone. Several lipid vitamin derivatives (*Chapter 10*), such as retinoic acid and 1α,25-dihydroxycholecalciferol, also show hormonal-like activities.

Class	Examples	Associated glands
Amines	Adrenaline, noradrenaline, thyroid hormones T_3 and T_4	Adrenal medulla, thyroid gland
Peptides and proteins	Insulin, growth hormone	Islets of Langerhans, anterior pituitary
Steroids	Cortisol, aldosterone, testosterone	Adrenal cortex

Table 7.1 A structural classification of hormones with selected examples. ⏎

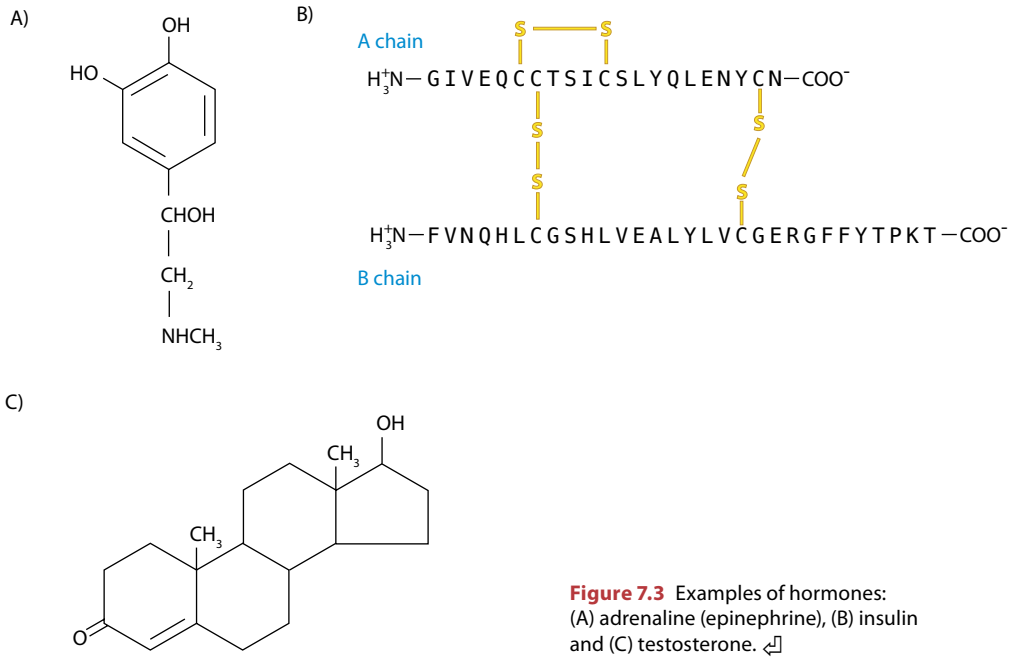


Figure 7.3 Examples of hormones: (A) adrenaline (epinephrine), (B) insulin and (C) testosterone.

7.2 HORMONE PRODUCTION

Amine hormones include adrenaline, noradrenaline (also called epinephrine and norepinephrine, respectively) and thyroid hormones. Their biosynthesis involves a series of enzyme-catalysed reactions in the cytoplasm of endocrine cells. For example, thyroid hormones (*Section 7.7*) are biosynthesised by iodination of tyrosine residues in the protein thyroglobulin found in the thyroid. Most peptide hormones are produced as large inactive prohormones, which are subsequently cleaved by enzymes to produce the active hormone. Sometimes several hormones may be derived from the same prohormone. Steroid hormones are biosynthesised by a sequence of enzymatic reactions using cholesterol as a common precursor. The enzymes responsible for conversion of cholesterol to hormone are located in the smooth endoplasmic reticulum and mitochondria of cells. The presence or absence of particular enzymes determines the type of steroid hormone produced by that specific cell.

Amine hormones, such as adrenaline and noradrenaline, are stored in secretory granules within the cytoplasm, but thyroid hormones are stored within the thyroid follicles as components of thyroglobulin. Peptide hormones are usually stored in membrane-bound vesicles within the cytoplasm of the endocrine cell. Steroid hormones are not usually stored but are released upon synthesis. However, the cytoplasm contains lipid droplets that contain precursor material for these hormones.

A number of factors control hormone production. The secretion of pituitary hormones is under the influence of peptides released from the hypothalamus and this, in turn, is influenced by signals from the central nervous system (CNS). Most hormones released from endocrine glands are controlled by a negative feedback effect, such as for thyroid hormones (*Section 7.7*) and cortisol (see *Figure 7.48*).

Hormones are released in response to nervous, hormonal or metabolic stimuli. Hormones stored in granules are released when the granules move and fuse with the plasma membrane. Some hormones, for example thyroxine, are released continuously, whereas others show diurnal variation; their release varies during the day. For example, cortisol concentrations are greater in the morning than at night. The concentrations of hormones in the plasma must be kept within narrow ranges for optimum function.

Changes in the amounts of materials regulated by hormones themselves may influence the release of that hormone, as is the case for insulin. Target cells and the liver contain enzymes that degrade hormones. Hormones of low M_r are removed from the circulation by the kidneys and excreted in urine.

The half-lives of hormones vary from a few seconds to weeks. Many small or water-insoluble hormones form complexes with large plasma transport proteins. The kidneys cannot remove these large complexes and so their rapid loss is prevented. In addition, these complexes protect the hormone from degradation by enzymes and release the hormone slowly. The bound and free hormones are in equilibrium and it is only the free fraction that is biologically active.

BOX 7.1 THE PINEAL GLAND

The pineal gland or body is an appendage of the posterior end of the roof of the third ventricle in the brain (*Figure 7.4*). It is shaped like a pine cone (hence its name), weighs 100–150 mg and is only 8–10 mm long. It is composed largely of pinealocytes and glial cells. In older individuals, the pineal gland often contains calcium deposits, sometimes referred to as brain sand! René Descartes (1596–1650) regarded the pineal gland as 'the seat of the soul' because he believed it was unique in the human brain in being the only structure not duplicated on the right and left sides. This observation is not strictly true, however, since it is finely divided into two hemispheres that can be observed by microscopy.

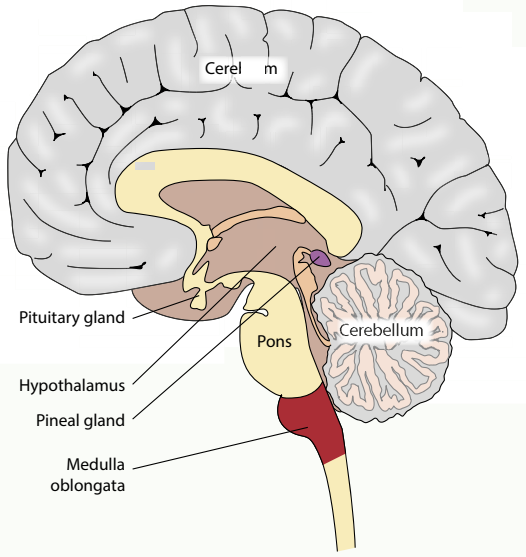


Figure 7.4 Schematic showing the location and structure of the pineal gland.

The physiological function of the pineal gland in humans is poorly understood. However, pinealocytes are known to produce the hormone melatonin (5-methoxy-*N*-acetyltryptamine) from tryptophan by acetylating serotonin (*Figure 7.5*). Melatonin has a relatively simple structure, yet communicates information about environmental lighting to physiological systems of the body. This light-transducing ability has led some to call the pineal gland the 'third eye'. Melatonin, therefore, helps regulate the internal 'clock' of the body, affecting biological rhythms and, in particular, patterns of sleeping and waking and possibly the onset of puberty. It may also function as a free radical scavenger and reduce oxidative damage. Most melatonin is synthesised during the night and can be secreted directly into the blood because the capillaries of the pineal gland are permeable and do not form part of the normal blood–brain barrier (see *Figure 3.8*).

In a normal environment in healthy humans, the release of melatonin usually starts at 21:00–22:00 and ends between 07:00 and 09:00, with peak levels of 60–70 pg/mL occurring in the plasma of adults between 02:00 and 04:00. The activity of the usual rate-limiting enzyme, serotonin-*N*-acetyltransferase (NAT), increases 7- to 150-fold during peak production. In daylight, the activity of NAT, and therefore melatonin production, is reduced. The rhythm is endogenously synchronised to 24 hours by the suprachiasmatic nucleus of the hypothalamus, an area of the brain known to coordinate biological clock signals, but can be entrained to values between 23 and 27 hours, primarily by the environmental light–dark cycle acting through the retina. When the retina is exposed to light, nervous impulses are relayed to the suprachiasmatic nucleus.

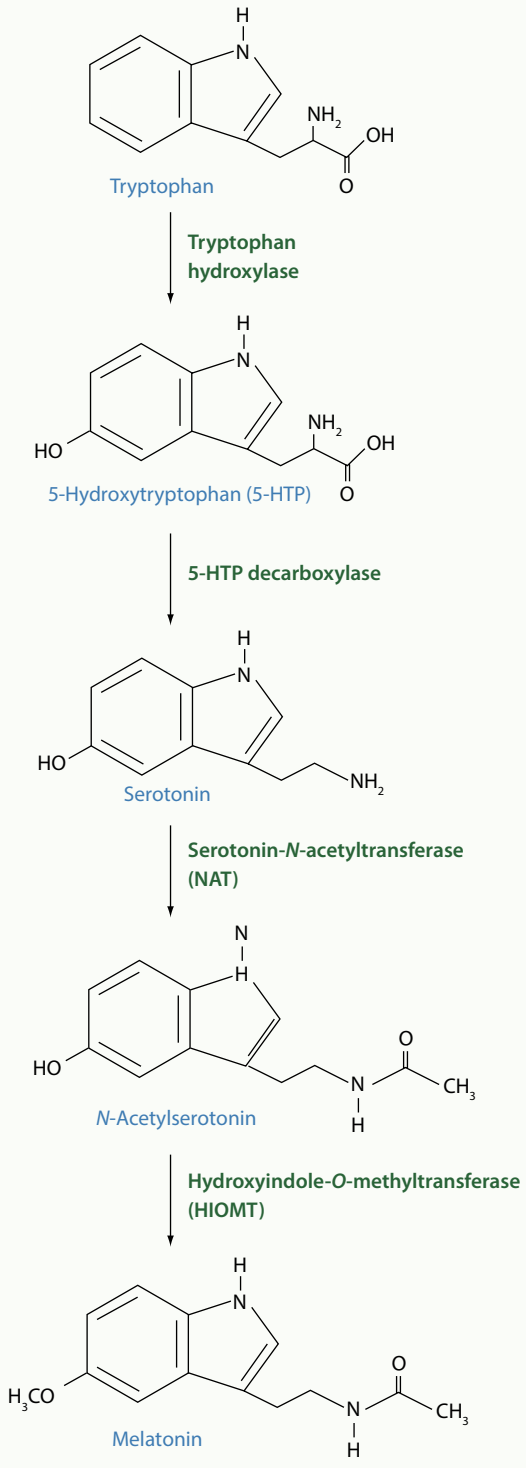


Figure 7.5 The biosynthesis of melatonin.

Nerve fibres from the hypothalamus descend to the spinal cord, from where postganglionic fibres ascend back to the pineal gland. Thus, the pineal gland can measure day length and adjust its secretion of melatonin appropriately.

Small cysts are commonly seen within the pineal region, although their discovery is frequently incidental to radiographic investigations. However, they are benign and non-progressive and should be treated conservatively. The pineal gland is also subject to numerous types of malignant tumours (*Chapter 17*), for example teratomas, germinomas, choriocarcinomas, endodermal sinus tumours, mixed germ cell tumours, pineoblastoma and pineocytoma, and gliomas. Fortunately, all are rare and collectively account for less than 1% of intracranial space-occupying lesions. With the exception of parenchymal cell tumours, pineoblastoma and pineocytoma, they occur mainly in patients below the age of 20 years. Germinomas and teratomas occur predominantly in males. The most common symptoms are secondary to hydrocephalus, such as headaches, vomiting and drowsiness, with visual problems, diabetes insipidus (*Section 7.8*) and reproductive abnormalities. In children, pineal tumours of the region are often associated with abnormal pubertal development. Some evidence suggests that precocious puberty is due to the production of human chorionic gonadotrophin (hCG) by germ cell tumours of the pineal gland. Delayed puberty has also been associated with pineal tumours.

Pineal tumours can be diagnosed from the symptoms, physical examination and tests for tumour markers in samples of cerebrospinal fluid (CSF), such as α fetoprotein (*Figure 7.6*) and hCG, together with CSF cytology. Computed tomography (CT) or magnetic resonance imaging (MRI; *Chapters 13* and *18*) can aid in differential diagnosis. Germinomas respond well to radiation therapy, whereas surgery is usually the choice for other types. There have been significant improvements in prognosis, with 62% surviving germinomas for over five years, although survival is only 14% for other malignant tumours.

Given that melatonin affects normal sleep patterns, it is used sometimes in combination with **phototherapy** to treat sleep disorders. Phototherapy, more properly called full spectrum bright light therapy, is repeated exposure to three to six hours of artificial light that is as bright as sunlight. The use of melatonin in shift workers, who often find it difficult to adjust to working at night and sleeping during the day, unfortunately has not shown promise. Melatonin therapy does not appear to help their condition and appears not to be as effective as phototherapy. However, melatonin therapy appears to be modestly beneficial in elderly insomniacs, who have lower concentrations than matched noninsomniacs. Jet lag also involves the disruption of circadian rhythms, where melatonin therapy seems to be of use. During long-distance air travel, taking melatonin close to the bedtime of the destination appears to reduce the symptoms associated with jet lag. The greatest benefits occur when jet lag would be expected to be greatest, that is, on journeys that cross many time zones.

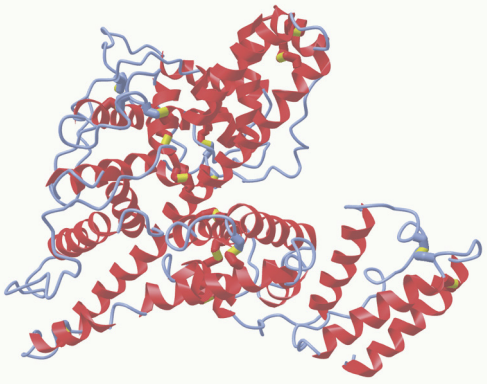


Figure 7.6 Molecular model of α fetoprotein. PDB files 7YIM.

Melatonin profiles show a seasonal phase, with an earlier secretion in summer than in winter. The condition **seasonal affective disorder** (SAD) or *winter depression* is characterised by changes of mood and eating and sleeping patterns. It appears to develop in people living at high latitudes in winter when sunlight is lacking. The increased release of melatonin in winter has been suggested as a possible cause of SAD but at present there is no consensus on its causes. The assumption that melatonin duration is a seasonal signal in humans led to the treatment of SAD with bright light in an attempt to induce advancement of the melatonin rhythm. It appears to be somewhat more efficient when given in the morning, although there is a large placebo effect. However, other mechanisms are also possible; indeed, many pharmacological antidepressant treatments stimulate melatonin secretion.

7.3 MECHANISMS OF HORMONAL ACTION

Hormones act by binding to specific receptors of target cells to form a complex (*Figure 7.7*) that initiates intracellular events, which produce the biochemical and physiological effects associated with the hormone. Only a target tissue will express the receptor for a given hormone and thus be able to respond to it.

Hormone receptors and second messengers

Hormone receptors may be located on the surface of the cell or within the cell.

The two major types of cell surface receptors are:

1. G protein-coupled receptors
2. receptor tyrosine kinases

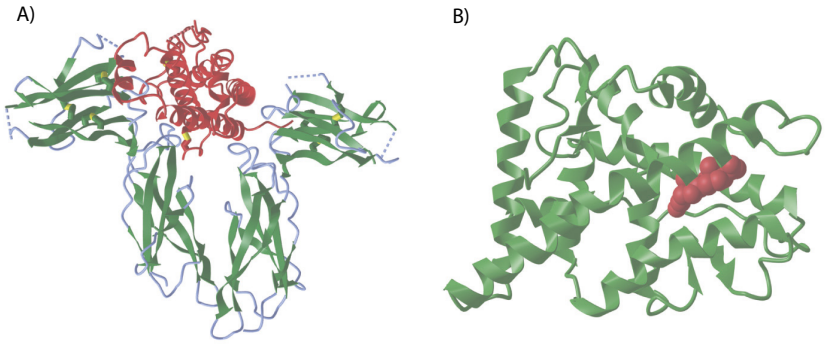


Figure 7.7 Molecular models of (A) human growth hormone bound to the extracellular portion of its receptor and (B) the steroid hormone progesterone bound to its intracellular receptor. In each case, the hormone is shown in red and its receptor in green. PDB files 3HHR and 1A28.

G protein-coupled receptors

G protein-coupled receptors (GPCRs) are integral trans-plasma membrane proteins that contain seven transmembrane α helices (*Figure 7.9*). Specific loops between the helices form extracellular and intracellular binding sites for the hormone and a G protein, respectively.

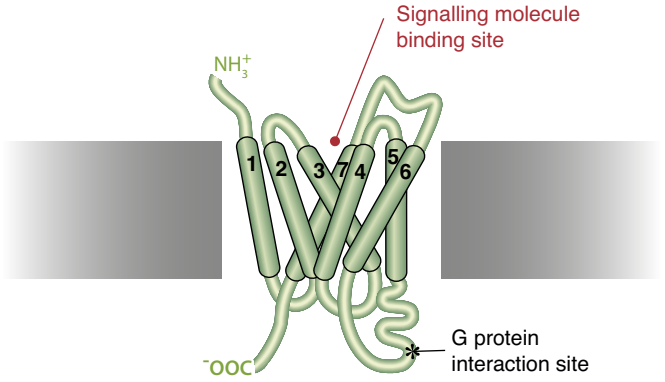


Figure 7.9 Schematic representation of a G protein-coupled receptor (GPCR).

Protein, peptide and amine hormones are water soluble and cannot easily cross the lipid bilayer of the plasma membrane. These types of hormones, such as glucagon, vasopressin, adrenaline and thyroid-stimulating hormone (TSH) (also most cytokines, growth factors and prostaglandins [*Chapter 10*]), bind to GPCRs and act through what are called second messenger systems, the hormone being the first or primary messenger. Second messengers are small M_r water-soluble ions or molecules that are generally able to move freely throughout the cell. The most common secondary messengers (*Figure 7.10*) are cyclic adenosine monophosphate (cAMP), the structurally related cyclic guanosine monophosphate (cGMP), Ca^{2+}, inositol triphosphate (IP_3) and diacylglycerol (DAG).

The binding of the hormone to the extracellular portion of the GPCR changes its conformation (*Figure 7.11*), such that the intracellular part can then activate a specific G protein. The G protein may, in turn, activate adenylate kinase, which catalyses the conversion of ATP to the secondary messenger cAMP:

$$ATP + H_2O \rightarrow cAMP + PP_i$$

The concentration of cAMP in the cytoplasm therefore increases and stimulates a protein kinase, which then catalyses the phosphorylation of specific enzymes in the cytosol. Depending on the kinase involved, phosphorylation can cause an increase or decrease in the activities of these cytosolic enzymes. The binding of the hormone has therefore initiated a signal along a transduction pathway consisting of several enzymes. The signal, however, is not only propagated but amplified, since the molecule at any one step is able to activate many subsequent ones in a chain reaction. Phosphodiesterase (PDE) activity inactivates cAMP by hydrolysing it to AMP:

$$cAMP + H_2O \rightarrow AMP$$

This prevents cAMP accumulating in the cytoplasm.

Margin Note 7.1 G proteins

G proteins are so called because they bind guanosine nucleotides (GDP and GTP). They are a diverse group, ranging from small soluble monomers to large multisubunit membrane proteins. However, all are GTPases that can autohydrolyse their bound GTP to GDP:

$$GTP\text{-protein} + H_2O \rightarrow GDP\text{-protein} + P_i$$

The green and red colours indicate the active (functional) and inactive conformations of the G protein (*Figure 7.8*). Thus, G proteins provide on–off switches in biochemical and physiological systems, as they oscillate between the two forms. They have wide ranging functions: from activities associated with numerous hormones to sense perception and the nucleocytoplasmic transport described in *Chapter 16*.

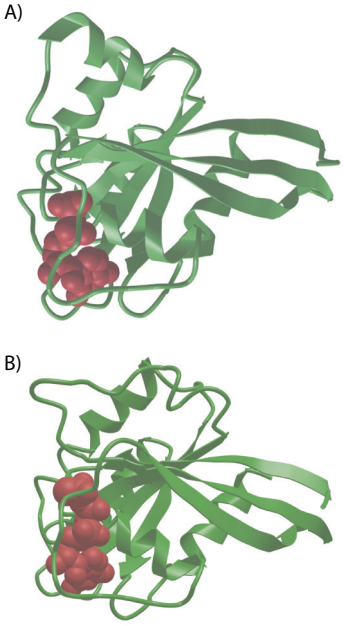


Figure 7.8 The small soluble G protein Rab11a with (A) a bound GDP and (B) a GTP. In both cases, the nucleotide is shown in red. Note the difference in conformation between the two forms. PDB files 1OIV and 1OIW, respectively.

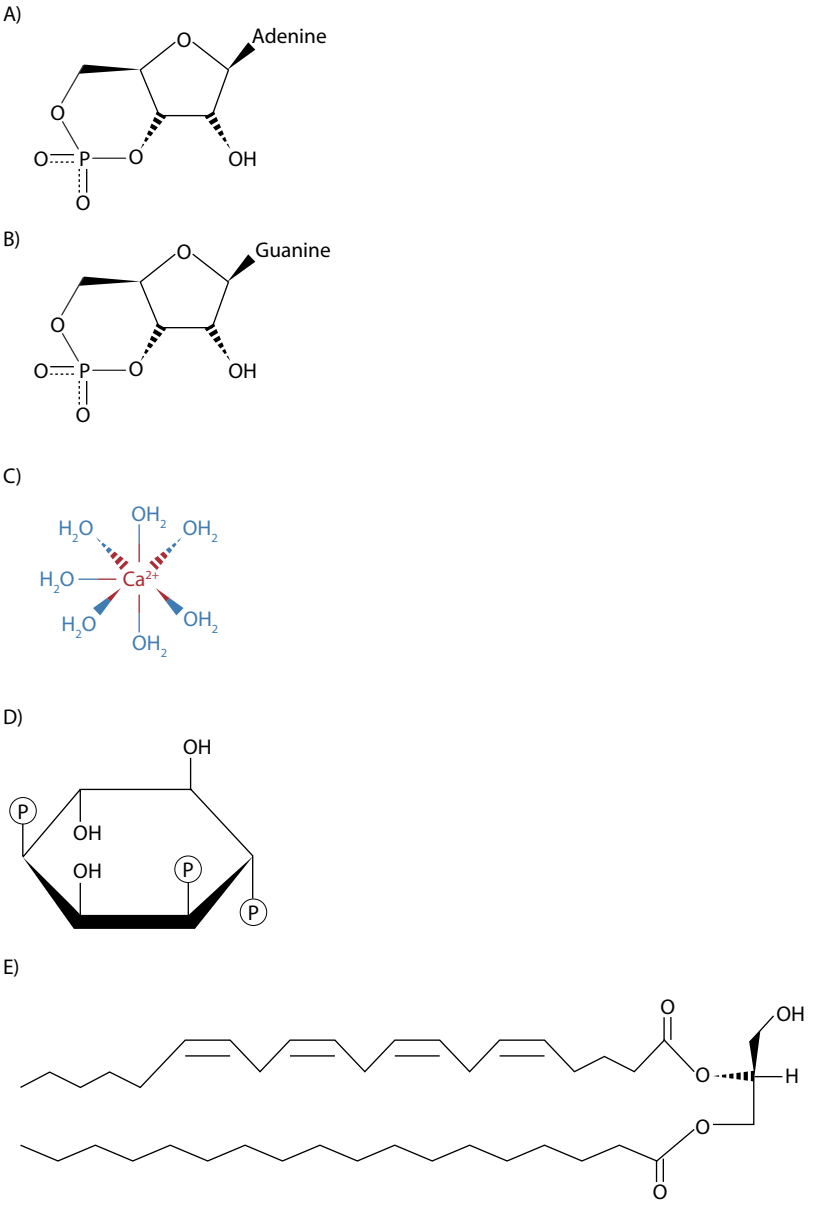


Figure 7.10 Examples of hormone secondary messengers. (A) Cyclic adenosine monophosphate (cAMP), (B) cyclic guanosine monophosphate (cGMP), (C) a hydrated Ca^{2+}, (D) inositol triphosphate (IP_3) and (E) a diacylglycerol (DAG). ⏎

Receptor tyrosine kinases

Receptor tyrosine kinases (RTKs), like GPCRs, have extracellular sites for hormone binding. They also have a cytoplasmic domain that is a potential protein kinase. Hormones that bind to RTKs include insulin and growth hormone. However, other signalling molecules (such as many growth factors, for example epidermal growth factor, and cytokines like erythropoietin [see *Figure 8.2*]) also act through RTKs. The binding of a hormone to its RTKs usually results in two receptor–hormone complexes associating to form a dimer, although the insulin receptor is an exception to this dimerisation rule. Following dimerisation, or with insulin its binding to its receptor, conformational changes in the complex activate the protein kinase domain, which autocatalyses the transfer of a variable number of phosphates from ATP to tyrosine residues of the receptor (*Figure 7.12*). This, in turn, activates a relay of kinases in the cytoplasm that propagates and amplifies the initial hormone signal. One hormone–receptor dimer can initiate several different intracellular signalling pathways simultaneously because various combinations of phosphorylated tyrosine residues lead to the activation of different relay kinases. This ability to trigger many pathways is a major difference between the actions of RTKs and GPCRs.

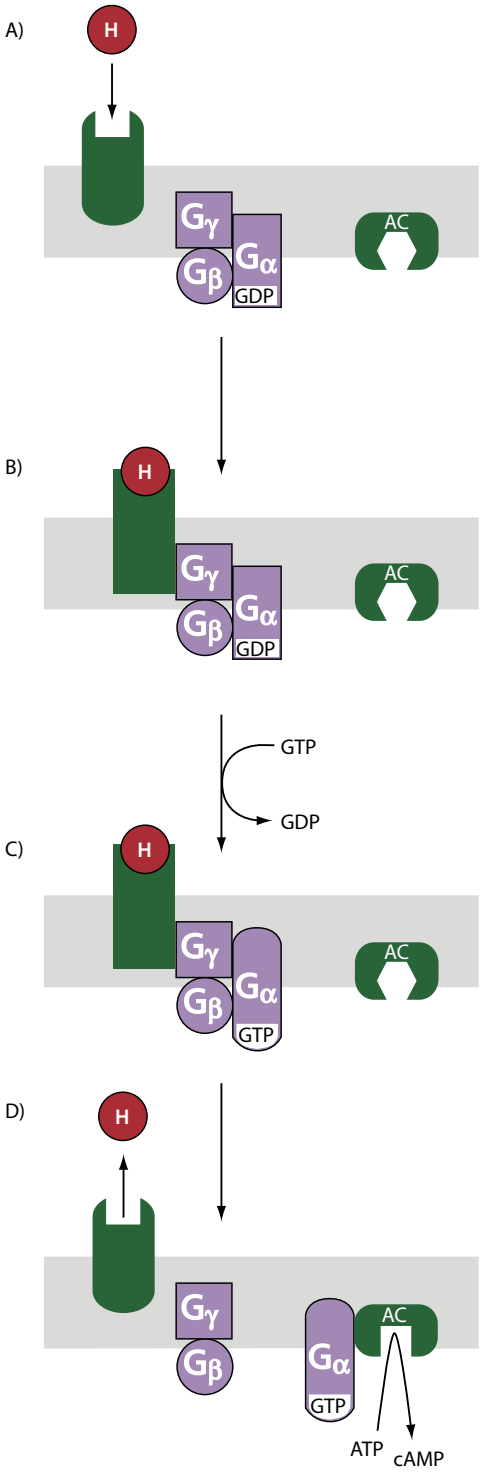


Figure 7.11 Mechanism of action of a hormone that binds to a G protein-coupled receptor (GPCR) and acts via the second messenger cAMP. Note the conformational changes that are shown schematically as the hormone binds to (A) its receptor, (B) and (C) the trimeric G protein as GDP and GTP are exchanged, and (D) the adenylate cyclase (AC) as it is activated by the Gα–GTP complex. ⏎

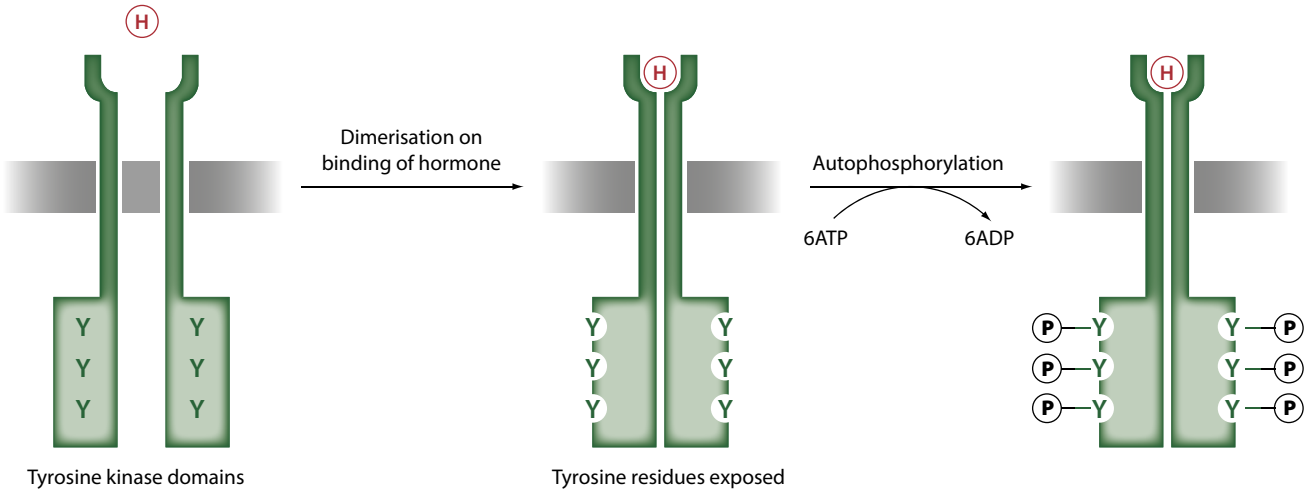


Figure 7.12 Schematic illustrating hormone binding, dimerisation and activation of a receptor tyrosine kinase (RTK). See also *Figure 7.7.*

Intracellular receptors

Receptors for steroid hormones and lipid vitamin derivatives, such as retinoids and 1α,25-dihydroxycholecalciferol (*Chapter 10*), are located in the cytosol or nucleus. Thus, these signalling molecules must enter their target cells to bind to their receptors and stimulate appropriate biochemical and physiological effects.

Steroid and thyroid hormones and retinoid signalling pathways

Steroid hormones are extremely hydrophobic and are transported by the blood to target cells bound to specific carrier proteins (*Figure 7.13*) or serum albumin. Their release at target cells allows them to diffuse directly across the plasma membrane into the cytoplasm.

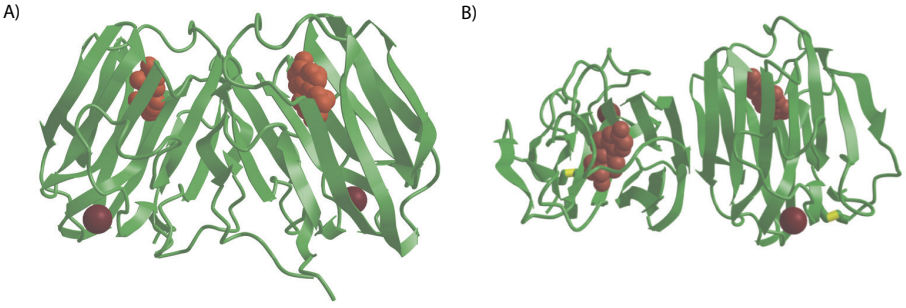


Figure 7.13 Molecular models of (A) dihydrotestosterone, the active form of testosterone, and (B) oestradiol, both bound to sex hormone binding globulin (SHBG). In both cases, the hormones are coloured red. PDB files 1KDM and 1LHU, respectively.

In the cytoplasm, their fate depends upon the nature of the hormone and its receptor. For example, glucocorticoid receptors are normally found in the cytosol as part of a multiprotein complex with heat shock proteins (HSPs; *Margin Note 7.2*). Thus, the binding of a glucocorticoid, for example cortisol, to its receptor induces a conformational change that releases the HSPs and exposes a nuclear locating signal (NLS; *Chapter 16*), which allows the hormone–receptor complex to be translocated into the nucleus. Here, the hormone–receptor binds to DNA at specific transcription regulatory sites called glucocorticoid response elements (GREs) that are part of a generalised group called hormone response elements (HREs), which respond to steroid and thyroid hormones and retinoid–receptor complexes. The binding of the first complex facilitates the attachment of a second hormone–receptor to the same GRE (*Figure 7.14*). The usual response to binding of a dimer is an increase in transcription of specific genes leading to the

Margin Note 7.2
Heat shock proteins

Heat shock proteins (HSPs) are a family of ubiquitous (found in eucarya, bacteria and eukaryotic organisms) proteins produced by cells in response to stress. They were first described in bacteria following their exposure to heat stress, hence the name. However, they are produced following many types of biological stress. Heat shock proteins participate in a range of cellular functions, including the folding of newly synthesised polypeptides, the refolding of metastable proteins, the degradation of misfolded proteins and the assembly, stabilisation and dissociating of protein complexes.

production of mRNA molecules, which are translated, giving rise to new proteins in response to the hormone signal. However, in some cases, the dimer complex inhibits transcription by binding to an inhibitory GRE that recruits a histone deacetylase, whose activity promotes the tight coiling of chromatin (termed condensation), which prevents transcription.

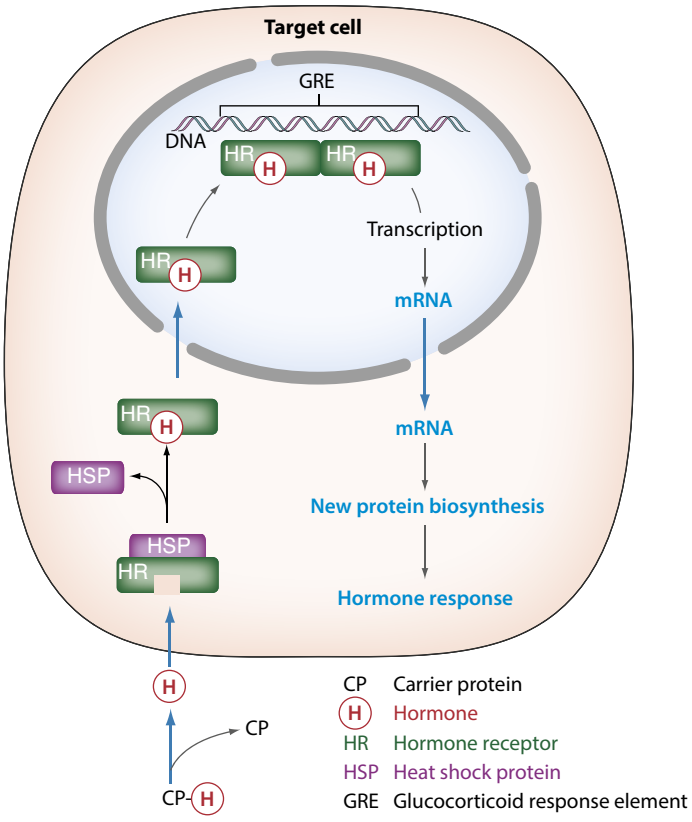


Figure 7.14 Schematic illustrating a steroid hormone (glucocorticoid) signalling pathway.

The receptors for thyroid hormones and retinoic acid have NLSs and are mainly found in the nucleus. On entering the cytoplasm, these types of signalling molecules migrate into the nucleus where they bind to their specific receptors and displace the HSP from the complex. The hormone–receptor is now able to bind to its appropriate response element and mediate transcription.

The desired physiological responses to steroid and thyroid hormones and retinoids may take hours to several days to become apparent. All mediate their effects by altering the transcription of specific genes. Hence, the response time reflects the interval between the signal being produced and the production of the necessary proteins in the target cell (*Figure 7.14*).

7.4 CAUSES OF ENDOCRINE DISORDERS

Endocrine disorders arise because a disruption to the endocrine system causes decreased (hypofunction) or increased (hyperfunction) hormonal activity or because of resistance to hormone action. There may be defects in the biosynthesis of the hormone due to an inherited deficiency of an enzyme required for its synthesis. Inappropriate stimuli may impair the release of the hormone or certain drugs may stimulate hormone release. Defects in negative feedback mechanisms may cause abnormal hormone secretion. Faulty inactivation or excretion of hormones in liver or renal diseases, respectively, can increase hormone concentrations. Excessive hormone secretion can occur ectopically from a nonendocrine source, such as a tumour. Even if correctly biosynthesised and released into the circulation, the target tissues may not recognise the hormone because of a lack of receptors or because the receptors themselves are nonfunctional. This is referred to as hormone resistance. Disorders will also occur if the target cells do recognise the hormone but there is a defect in the secondary messenger system responsible for converting the hormonal signal to a biochemical or physiological action.

In some autoimmune diseases (*Chapter 5*), antibodies are produced that stimulate or destroy endocrine glands, as in Graves disease and autoimmune thyroiditis, respectively.

The various causes underlying endocrine disorders are summarised in *Figure 7.15*. A considerable number of endocrine disorders have been described, including disorders associated with the pituitary, thyroid, pancreas and adrenal glands and the reproductive systems.

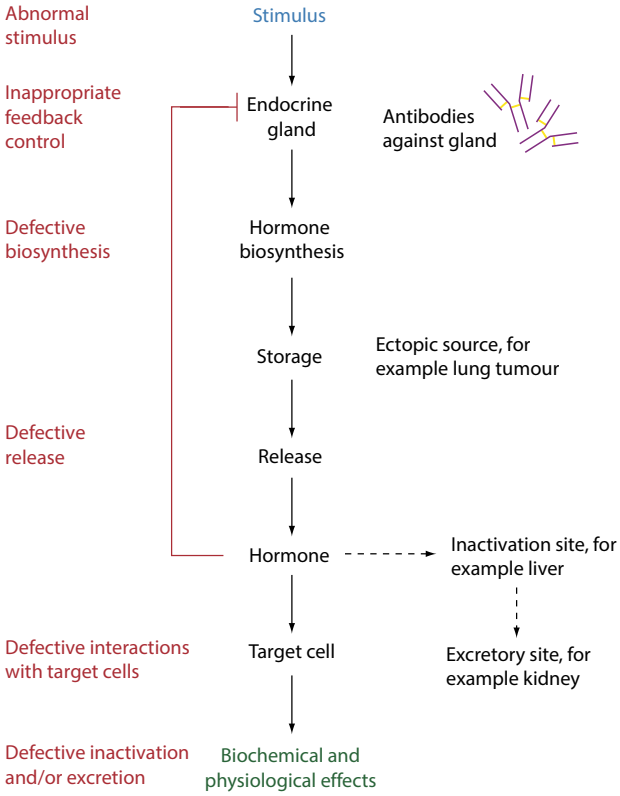


Figure 7.15 Overview of the causes of endocrine disorders.

7.5 THE PITUITARY GLAND

The pituitary gland is often referred to as 'the master gland', given that its secretions regulate the activities of many of the other hormone-producing glands. Despite its crucial role, it weighs only about 0.5 g. The pituitary gland is found in a bony cavity at the base of the skull and is connected to the hypothalamus by a pituitary stalk composed of blood vessels and nerve fibres. It has two lobes, an anterior pituitary or adenohypophysis and a posterior pituitary or neurohypophysis (*Figure 7.16*).

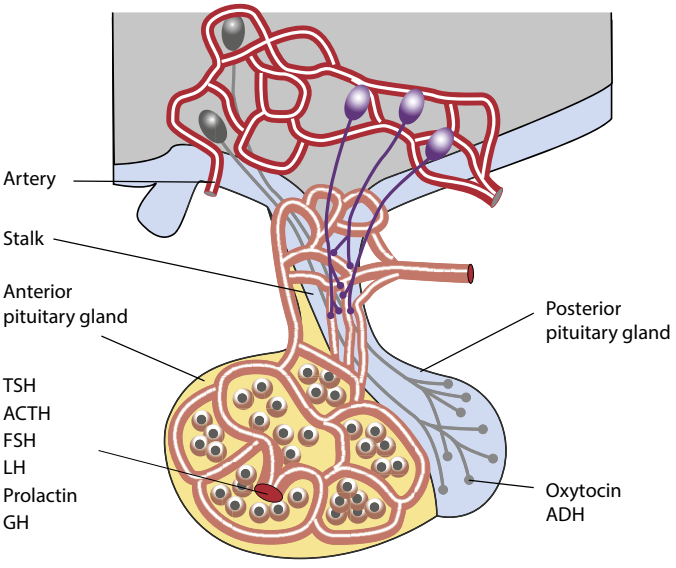


Figure 7.16 Structure of the pituitary gland and the hormones it secretes. ACTH, adrenocorticotrophic hormone; ADH, antidiuretic hormone; FSH, follicle-stimulating hormone; GH, growth hormone; LH, luteinising hormone; TSH, thyroid-stimulating hormone.

The anterior pituitary secretes a number of hormones (*Figure 7.17*) that are regulated by the release of peptides from the hypothalamus with stimulatory or inhibitory effects on the anterior pituitary. The principal peptides released by the hypothalamus reach the anterior pituitary by a portal blood circulation. They include thyrotrophin-releasing hormone (TRH), growth hormone-releasing hormone (GHRH), gonadotrophin-releasing hormone (GnRH), corticotrophin-releasing hormone (CRH) and dopamine. Apart from dopamine (see *Figure 7.23*), all are stimulatory, controlling the release of thyroid-stimulating hormone (TSH), growth hormone (GH), follicle-stimulating hormone (FSH), luteinising hormone (LH) and adrenocorticotrophic hormone (ACTH).

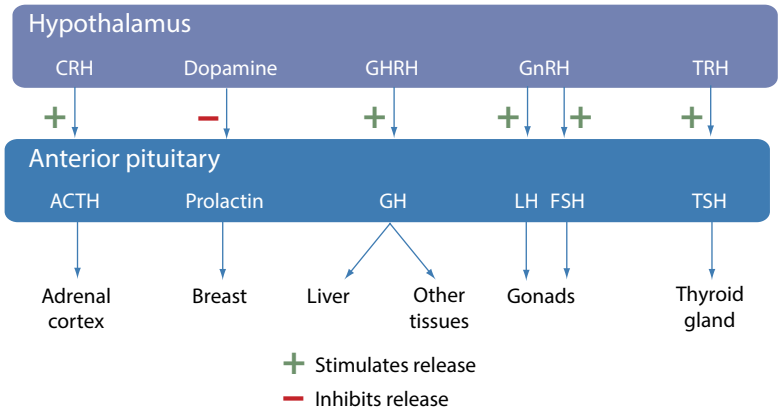


Figure 7.17 Hypothalamic factors that regulate anterior pituitary function. ACTH, adrenocorticotrophic hormone; CRH, corticotrophin-releasing hormone; FSH, follicle-stimulating hormone; GH, growth hormone; GHRH, growth hormone-releasing hormone; GnRH, gonadotrophin-releasing hormone; LH, luteinising hormone; TRH, thyrotrophin-releasing hormone; TSH, thyroid-stimulating hormone. ⏎

Thyroid-stimulating hormone (*Figure 7.18*) targets the thyroid gland causing it to release thyroid hormones.

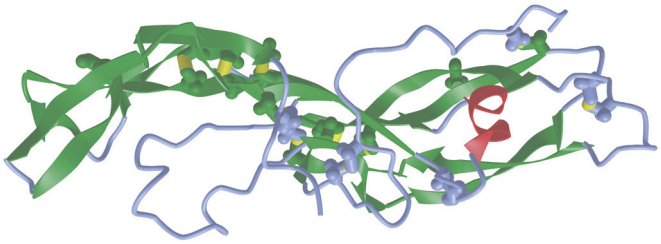


Figure 7.18 Molecular model of thyroid-stimulating hormone (TSH). PDB files 7T9I. ⏎

Growth hormone (see *Figures 7.7* and *7.25*) acts on general body tissues and promotes growth and development. In contrast, gonadotrophins, such as LH and FSH (*Figure 7.19*), act on the testes and ovaries and are essential for reproduction. Adrenocorticotrophic hormone acts on the adrenal cortex and stimulates the release of cortisol. Prolactin stimulates the female mammary glands, initiating lactation. The release of prolactin is inhibited by dopamine.

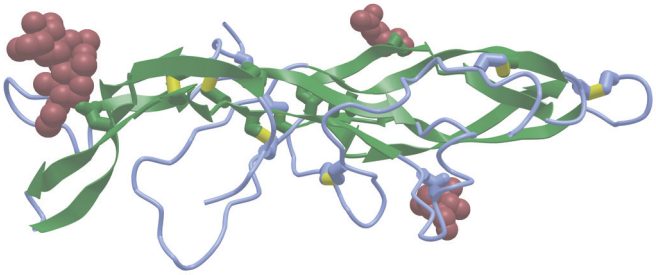


Figure 7.19 Molecular model of follicle-stimulating hormone (FSH) PDB file 1FL7. ⏎

The posterior pituitary is composed of a collection of nerve fibres originating in the hypothalamus. It secretes two main hormones, antidiuretic hormone (ADH)/vasopressin and oxytocin (*Figure 7.20*), although both are made in the hypothalamus and reach the posterior pituitary in the nerve fibres of the stalk. The former stimulates the kidneys to conserve water (*Chapter 8*), whereas the latter promotes uterine contractions during childbirth and stimulates the release of milk in breastfeeding.

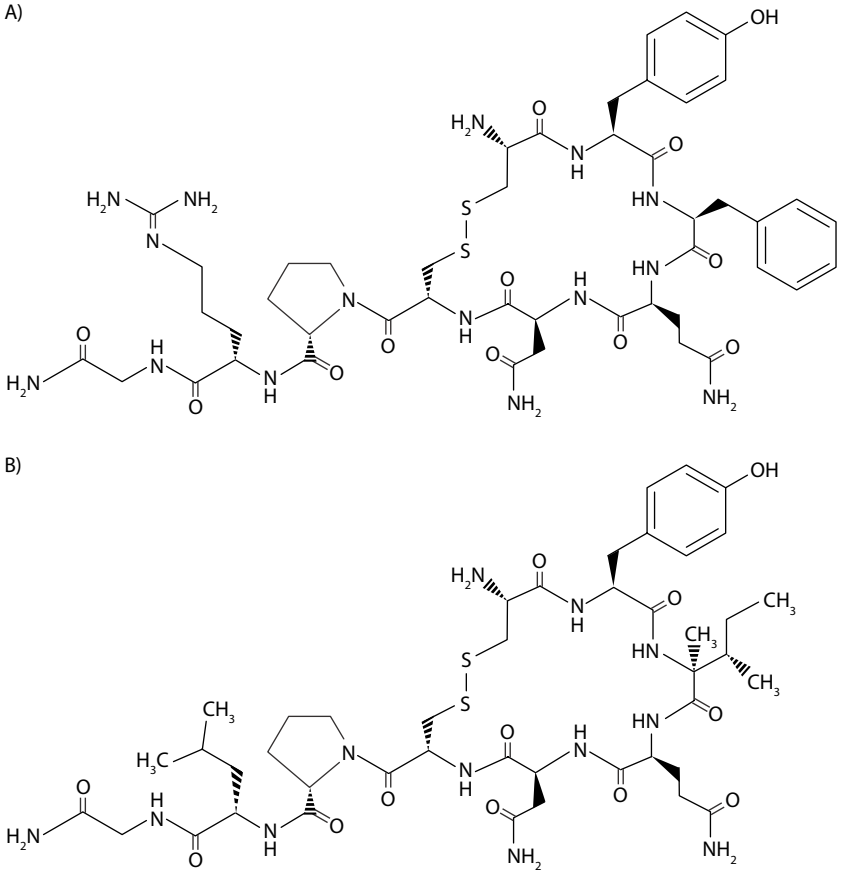


Figure 7.20 (A) Antidiuretic hormone (ADH)/vasopressin and (B) oxytocin.

Disorders of pituitary function

The majority of disorders in pituitary function are caused by tumours of the gland, although some pituitary diseases can lead to underproduction of its hormones. Approximately 80% of pituitary tumours are functional, that is, they cause an excessive secretion of hormones. The most common secrete prolactin (*Figure 7.21*) and account for 50% of functional tumours. The incidence of tumours affecting other hormones is 15% GH, 10% ACTH, 4% FSH/LH and less than 1% promoting TSH secretion. Large pituitary tumours may also exert pressure on nerves, causing headaches and visual disturbances. Radiological investigations such as X-rays, CT and MRI, outlined in *Chapter 18*, are useful methods of locating tumours and estimating their sizes.

Tumours secreting prolactin, prolactinomas, cause hyperprolactinaemia that, in turn, can lead to infertility in both males and females. Nonprolactin-secreting tumours or compression of the pituitary stalk, which block the dopaminergic inhibition of prolactin secretion, also result in hyperprolactinaemia. Hyperprolactinaemia abolishes menstruation (**amenorrhoea**) and causes an inappropriate release of breast milk (**galactorrhoea**), impotence and breast development (**gynaecomastia**) in males. Other causes of hyperprolactinaemia include drugs, for example phenothiazines, which block dopamine receptors, or methyldopa, which reduces the level of dopamine in the brain (*Figure 7.22*).

Investigating a patient with a possible prolactinoma includes assessing the concentrations of prolactin in the plasma (*Table 7.2*) following stimulation with TRH, although this test is not commonly used in most hospitals. In addition to TRH, prolactin is also

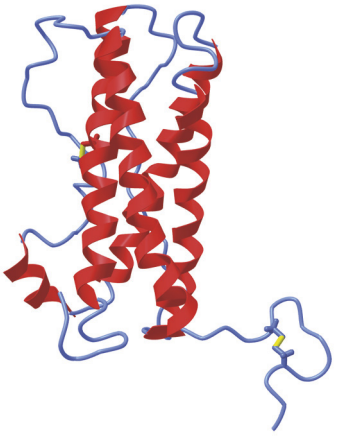


Figure 7.21 Molecular model of prolactin. PDB file 1RW5.

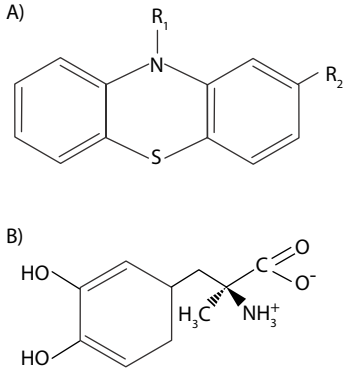


Figure 7.22 (A) The generalised structure of a phenothiazine and (B) the structure of methyldopa.

[Prolactin] (mU/L)	Interpretation
Males <381	Normal
Females <629	Normal
Males 381–1000	Repeat test; does not usually indicate a serious problem
Females 629–1000	Repeat test; does not usually indicate a serious problem
1000–2000	Repeat test; the increased [prolactin] may be secondary to stress, drug use, hypothalamic disorders, acromegaly, primary hyperthyroidism or chronic renal failure
2000–4000	Possible microprolactinoma or a hypothalamic disorder
4000–6000	Prolactinoma probably present, although the possibilities of a hypothalamic disorder or pregnancy should be investigated
>6000	Virtually always indicates the presence of a macroprolactinoma

Table 7.2 Interpretation of serum prolactin concentrations in the diagnosis of hyperprolactinaemia.

secreted in response to stress and oestrogens. Patients with a prolactinoma have plasma prolactin concentrations in excess of 2000 mU/L. These high values are generally not affected by TRH stimulation in individuals with prolactinomas. The first line of treatment is with a dopamine antagonist, such as bromocriptine (*Figure 7.23*), although surgical removal of the tumour may be necessary in cases that do not respond to drug therapy.

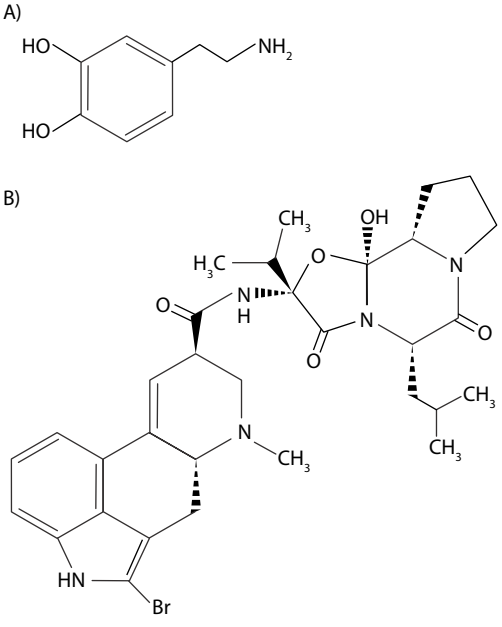


Figure 7.23 (A) Dopamine and (B) bromocriptine.

Antidiuretic hormone secretion is essential for life and disorders of its release are well recognised. The release of ADH is stimulated by increased osmolality of the plasma and a decrease in blood volume detected by hypothalamic osmoreceptors and cardiac baroreceptors, respectively. The role of ADH in fluid regulation is outlined in *Chapter 8*. A decreased output of ADH gives rise to diabetes insipidus, characterised by excessive production of dilute urine (polyuria). Patients are constantly thirsty (**polydipsia**), have hypernatraemia and present with a plasma osmolality in excess of 295 mmol/kg.

Cranial diabetes insipidus may be caused by brain tumours, meningitis, trauma and surgery, but can be **idiopathic**. Nephrogenic diabetes insipidus occurs when the kidneys fail to respond to ADH. The lack of response may be caused by drugs, such as lithium, chronic renal disease or hypercalcaemia (*Chapter 10*) or it may be congenital. Patients

with suspected diabetes insipidus are investigated by performing a fluid deprivation test: the patient is deprived of fluid intake for eight hours. In normal individuals, this results in concentrated urine with a plasma osmolality below 295 mmol/kg. However, in patients with diabetes insipidus, urine is concentrated and the plasma osmolality increases. Following the eight-hour period, the patient is allowed to drink water and given desmopressin (see *Figure 13.47*), a synthetic analogue of ADH, after which the urine is collected hourly for a further four hours. In cranial diabetes insipidus, the urine becomes concentrated, but with nephrogenic diabetes insipidus this does not occur, as the kidneys are insensitive to ADH (or desmopressin in this case). Hence, the test discriminates between cranial and nephrogenic diabetes insipidus.

Patients suffering from diabetes insipidus require access to rehydrating fluids but, in each case, the underlying cause must be treated. Patients with cranial diabetes insipidus are often given desmopressin in a nasal spray or chlorpropamide (*Figure 7.24*), which increases renal sensitivity to ADH. The use of the latter drug requires careful monitoring, since it can lead to hypoglycaemia. Individuals with nephrogenic diabetes insipidus do not respond to analogues of ADH and often there is no suitable treatment. Special attention to adequate fluid intake is essential.

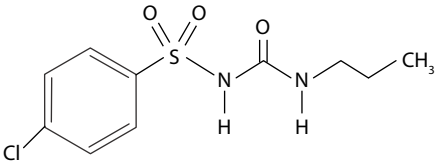


Figure 7.24 Computer drawn model of chlorpropamide.

Disorders of oxytocin are uncommon and have little clinical significance. A number of patients present with hypopituitarism, which is a failure to secrete one or more pituitary hormones, although this is a relatively uncommon complaint. Hypopituitarism may result from a tumour, infarction, infections or trauma affecting the pituitary or may be secondary to disorders of the hypothalamus. The clinical presentation of hypopituitarism often depends on the age of the affected individual. A decreased release of GH is often an early feature, leading to dwarfism in children (*Section 7.6*). Inadequate secretion of gonadotrophins may cause amenorrhoea (see above) and infertility in adult females, with the loss of secondary sexual characteristics in males. Elderly patients with hypopituitarism may present with hypoglycaemia and hypothermia, relating to ACTH and TSH deficiencies, respectively. In most cases, GH and gonadotrophin deficiencies tend to present before that of ACTH. Hyposecretion is assessed by stimulatory tests where the ability of the anterior pituitary to secrete the hormone in question is assessed following stimulation of the patient with the hypothalamic peptide or its analogue. A failure to respond would suggest hypopituitarism.

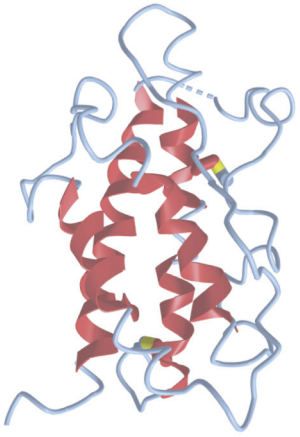


Figure 7.25 Molecular model of human growth hormone (GH). PDB file 1HGU.

7.6 GROWTH HORMONE DISORDERS

Growth hormone or somatotrophin (*Figure 7.25*; see also *Figure 7.7*) is a polypeptide 191 amino acid residues long. The essential function of GH is to promote linear growth and the maintenance of tissues. It also stimulates the uptake of amino acids and protein biosynthesis by cells, increases blood glucose concentration and fat metabolism, and promotes epiphyseal bone growth. These effects are mediated by locally acting effectors called somatomedins, which are biosynthesised by many tissues but particularly by hepatocytes. Somatomedins stimulate cell proliferation and/or their differentiation. They include insulin-like growth factors-1 and -2 (IGF-1 and IGF-2). Insulin-like growth factor-1 (*Figure 7.26*) is the most significant physiologically and, indeed, its concentration correlates with that of GH.

Approximately 70% of plasma GH is complexed with growth hormone binding protein. Growth hormone is biosynthesised in the anterior pituitary gland in an inactive form called preprogrowth hormone that is hydrolysed in several enzyme-catalysed reactions to give active GH prior to its secretion. Compare this with the secretion of insulin (see *Figure 7.41*). Several factors, such as sleep, amino acids, exercise and stress, stimulate GHRH release from the hypothalamus and that, in turn, stimulates GH secretion. Hyperglycaemia stimulates the secretion of somatostatin from the hypothalamus, and this inhibits the secretion of GH. Increasing concentrations of blood GH and IGF-1

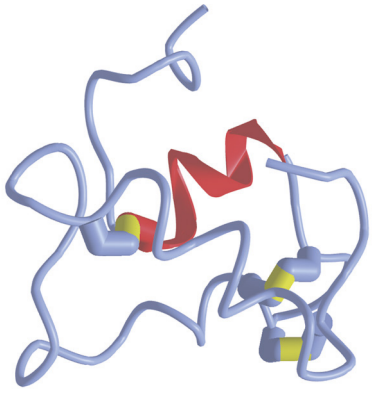


Figure 7.26 Molecular model of insulin-like growth factor-1 (IGF-1). PDB file 1B9G.

exert a negative feedback effect that prevents further release of GH (*Figure 7.27*). The rate of adult secretion varies but is generally about 1.4 mg daily and occurs in pulses, with the largest amounts being released during sleep.

The clinical features resulting from an excess or deficiency of GH depend on the age of the person. A deficiency during childhood leads to a stunted growth called dwarfism (*Figure 7.28*) and therefore requires early detection. However, GH deficiency is a rare cause of dwarfism, and other causes, for example thyroid deficiency or inadequate nutrition, must first be excluded. The most common causes of GH deficiency are non-endocrine tumours that affect the pituitary gland or hypothalamus. Growth hormone deficiency may also be a consequence of generalised pituitary disease or a congenital defect leading to a deficient production of GHRH. Whatever the cause, the major clinical feature is stunted growth well below that expected for a child of comparable age, with a short height and immature face and skeleton as revealed by radiological investigations. Clinical signs of other anterior pituitary hormone deficiencies may be evident.

Figure 7.28 Photographs showing one of the authors beside (A) a dwarf and (B) a person with combined gigantism and acromegaly.

The most common cause of excessive GH release is a GH-secreting pituitary tumour or adenoma (*Chapter 17*). Although these are benign, they are not subject to normal control and continually secrete large amounts of GH. The causes of these tumours are unknown, but a genetic basis has been suggested. Ectopic GH secretion is extremely rare but has been reported in patients with bronchial carcinoma.

Excessive GH release causes gigantism in children and acromegaly in adults (*Figure 7.28*). Children with excess GH grow as much as 150 mm annually, reaching seemingly abnormal heights, often more than 2.4 metres. Muscle weakness is seen in long-standing cases. Acromegaly has an insidious onset and may take years for its clinical features to become apparent. These include enlargement of bones of hands, feet and face, thickening of soft tissues causing coarse facial features, enlarged tongue and lips, prognathism or protruding jaw, increased sweating and enlargement of internal organs, such as liver, spleen and heart. Additionally, acromegalics suffer from paraesthesia of the hands and feet due to entrapment of nerves by thickened bone and subcutaneous tissue, headaches/vision disorders due to the growing pituitary tumour and sensory nerve entrapment, impaired glucose tolerance or diabetes mellitus, and increased incidences of coronary heart disease and stroke (*Chapter 14*). Individuals affected by excessive GH secretion throughout life show features of both gigantism and acromegaly. The prognosis for both gigantism and acromegaly depends on how far the disorder has advanced. Gigantism is rarely life threatening, and prognosis is usually good. However,

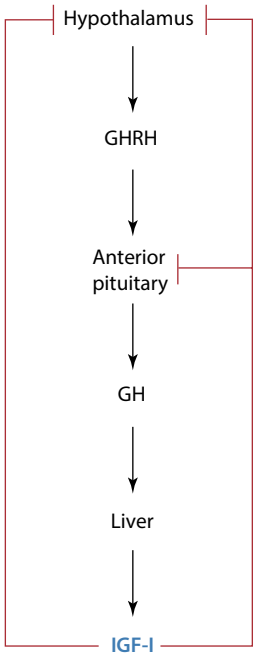


Figure 7.27 The regulation of human growth hormone secretion. GH, growth hormone; GHRH, growth hormone-releasing hormone; IGF-1, insulin-like growth factor-1.

an individual with advanced acromegaly will develop serious complications, such as coronary heart disease, cerebrovascular disease and diabetes mellitus.

Diagnosis and treatment of growth hormone disorders

A variety of tests are used to assess GH deficiency. Random measurements of serum GH are of limited value due to fluctuations in plasma GH concentrations in normal individuals. Urinary GH excretion is low when there is a deficiency, but obtaining an accurately timed collection of urine is difficult. Most tests rely on demonstrating that the hormone does not increase in concentration following a stimulus. Growth hormone increases after exercise and this has been used as a preliminary screening test. In the exercise test, the patient is subjected to hard physical exercise until they have a pulse rate greater than 150/min (*Chapter 14*). Blood is collected at 0, 2 and 20 minutes after stopping exercise. In normal individuals, the plasma concentration of GH increases by 20 mU/L above the initial value. Growth hormone release increases during sleep; hence, high values in a nocturnal sample may exclude deficiency. Blood samples are collected using a venous catheter at 30-minute intervals for three to four hours after the onset of sleep. A peak of at least 10 mU/L occurs in normal individuals but not in patients with GH deficiency. Clonidine (*Figure 7.29*) is a potent stimulator of GH secretion and is used as a definitive test for GH deficiency. Growth hormone from genetically engineered sources is used in treatment but must be continued until longitudinal growth is completed. When deficiency is due to low levels of GHRH, analogues of this peptide, for example hexarelin, have been used.

A diagnosis of excess GH is made on clinical grounds supported by biochemical and radiological investigations. Photographs taken of the patient when younger are particularly useful for diagnosis. Basal serum concentrations of GH are increased in GH excess but, because release is influenced by so many factors, the result of a single sample is unreliable.

The concentration of IGF-1 in serum is raised in patients with acromegaly and is of diagnostic significance. The oral glucose tolerance test (OGTT) is used to confirm a diagnosis of acromegaly and is similar to that used for diagnosis of diabetes mellitus (*Section 7.8*) except that concentrations of plasma GH are also determined. In a healthy individual, the glucose load suppresses GH release to below 2 mU/L by stimulating the release of somatostatin, which, in turn, inhibits GH release. In patients with acromegaly, this suppression is not seen.

The management of GH excess cannot reverse any clinical changes that have taken place prior to treatment. However, management is important because it improves survival and reduces deaths due to heart disease or stroke. Growth hormone-secreting tumours are treated in one of three ways, and the clinical decision depends on the severity of disease, age of the patient and response to treatment. Surgical removal of the pituitary tumour may be attempted, depending on its size and location, with steps taken to minimise damage to the anterior pituitary. If GH concentrations do not normalise, other treatments must be considered. Radiotherapy is usually performed over a four- to six-week period by external irradiation using a cobalt source. Its effects are slow and, in some cases, hyposecretion of other pituitary hormones may occur. Drugs, such as bromocriptine (see *Figure 7.23*), suppress GH release in acromegalics and are often used in patients too old to undergo surgery or radiotherapy. Somatostatin analogues, for example octreotide (*Figure 7.30*), have also been used to inhibit GH release. Some patients require all three modes of treatment but are not always successfully treated. To detect any recurrence, serum IGF-1 needs to be monitored at regular intervals.

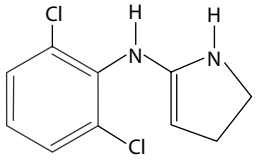


Figure 7.29 Clonidine.

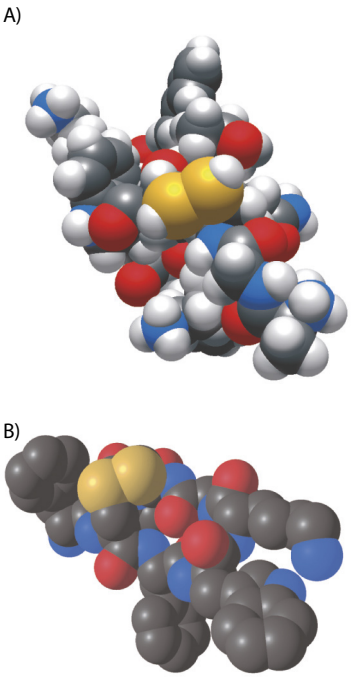


Figure 7.30 Molecular models of (A) somatostatin and (B) octreotide. PDB files 2MI1 and 6VC1, respectively.

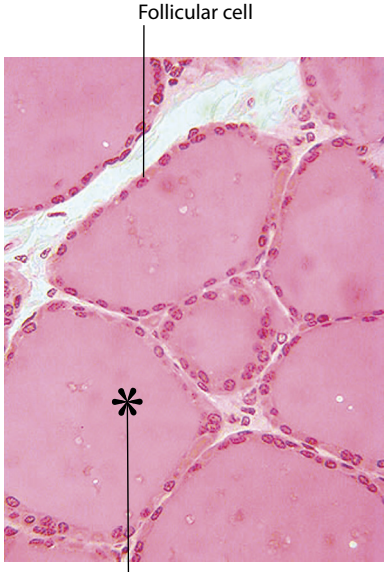


Figure 7.31 The histological structure of the thyroid gland. Courtesy of Dr A.L. Bell, University of New England, College of Osteopathic Medicine, USA.

7.7 THYROID HORMONE DISORDERS

The thyroid gland is the largest endocrine gland in the body, weighing about 20 g. Its bilobular structure consists of microscopic spherical follicles with secretory cells (*Figure 7.31*) that biosynthesise the thyroid hormones. Thyroid hormones (*Figure 7.32*) consist of triiodothyronine (T_3) and thyroxine (T_4). Triiodothyronine is four times more active than T_4 and has a half-life of 1.5 days compared with 9 days for T_4. However, in most tissues, particularly the liver, T_4 can be readily converted to T_3. More than 99% of T_3 and T_4 are transported in the serum as complexes: approximately 70% to thyroxine-binding globulin (TBG), about 20% to albumin and around 10% to prealbumin. The remaining small free portion is the metabolically active form.

The effects of thyroid hormones are to increase heat production and its associated oxygen consumption; stimulate the metabolism of proteins, fats and carbohydrates; and promote normal growth. They are also necessary for the normal functioning of the CNS. Plasma concentrations of T_4 and T_3 are regulated by the release of TSH from the

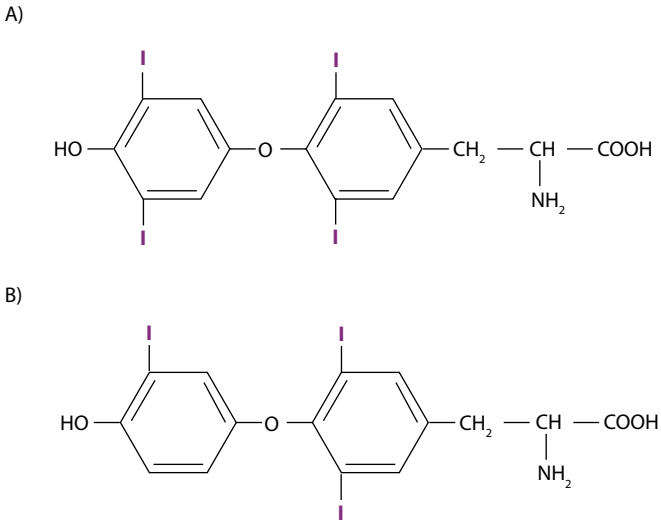


Figure 7.32 The major thyroid hormones: (A) triiodothyronine (T3) and (B) thyroxine (T_4).

anterior pituitary that, in turn, is controlled by the release of TRH from the hypothalamus. Increasing concentrations of free T_4 and T_3 inhibit further release of TSH and TRH by a negative feedback effect (*Figure 7.33*).

Overactivity of the thyroid gland and its associated excessive production of hormones can produce hyperthyroidism. The clinical syndrome resulting from hyperthyroidism is thyrotoxicosis (*Figure 7.34*), which presents with weight loss, sweating, heat intolerance, anxiety, hyperkinesis, increased appetite, osteoporosis, menorrhagia, tachycardia and pretibial oedema.

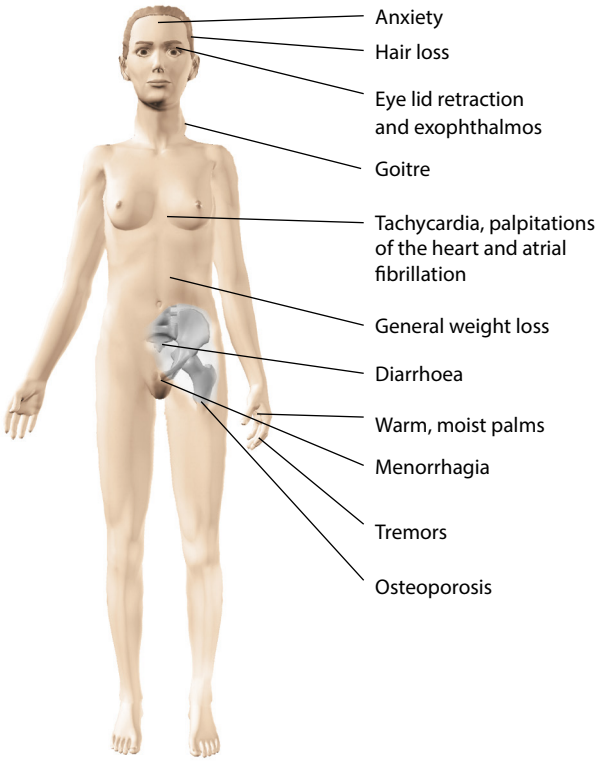


Figure 7.34 Schematic to show the clinical features of thyrotoxicosis.

The most common cause of hyperthyroidism is Graves disease, which can occur at any age but particularly in 20- to 40-year-old females. Patients with Graves disease suffer from **exophthalmos** or protrusion of the eyeballs, in addition to clinical features of hyperthyroidism. It is an autoimmune disease (*Chapter 5*), characterised by the presence of thyroid-stimulating antibodies in the blood that bind to TSH receptors in thyroid cells and stimulate them in a similar manner to TSH. Toxic nodules are the second main

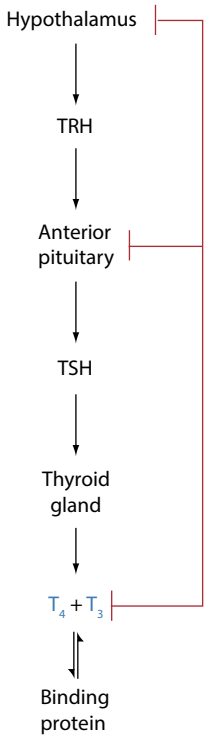


Figure 7.33 The negative feedback regulation of thyroxine (T_4) and triiodothyronine (T_3) secretion. TRH, thyrotrophin-releasing hormone; TSH, thyroid-stimulating hormone.

Margin Note 7.3
What s in a name?

Addison, Conn, Cushing and Graves diseases (*Sections 7.7* and *7.9*) are all named after the physicians who first described them:

- Thomas Addison (1793–1860), an English doctor
- Jerome Conn (1907–1994), an American endocrinologist
- Harvey Cushing (1869–1939), an American neurosurgeon
- Robert Graves (1796–1853), an Irish surgeon

cause of hyperthyroidism and tend to be found in elderly patients. They may occur singly or as multiples in a nodular goitre and are autonomous (self-governing) secretors of thyroid hormones. An excessive intake of T_4 in individuals who are treated for hypothyroidism can cause hyperthyroidism. Rare causes of hyperthyroidism include ectopic thyroid tissue and tumours that secrete TSH, although the latter are very uncommon.

Hypothyroidism (myxoedema) is most common in women of 30–60 years of age (*Figure 7.35*). Its clinical features include psychosis, diminished sweating, hypokinesis, weight gain, muscle weakness, constipation, dry and cold skin, dry hair, ischaemic heart disease, bradycardia and menstrual irregularities. Its primary cause is a defect in the secretion of thyroid hormones by the thyroid gland. In, for example, Hashimoto thyroiditis, an autoimmune disease, there is destruction of thyroid tissue by antibodies (*Chapters 4* and *5*) produced against the thyroid. Hypothyroidism may also occur after surgery (postthyroidectomy) and following treatment with antithyroid drugs. Congenital hypothyroidism occurs because of the failure of the thyroid gland to develop normally during embryonic growth. If untreated, this condition results in **cretinism** where the child suffers mental deterioration, muscle weakness, short stature and neurological signs and is often deaf and mute. Hypothyroidism may also arise due to iodine deficiency in particular parts of the world, for example certain regions of Africa. Secondary causes of hypothyroidism are linked to the pituitary or the hypothalamus with defective secretions of TSH and TRH, respectively.

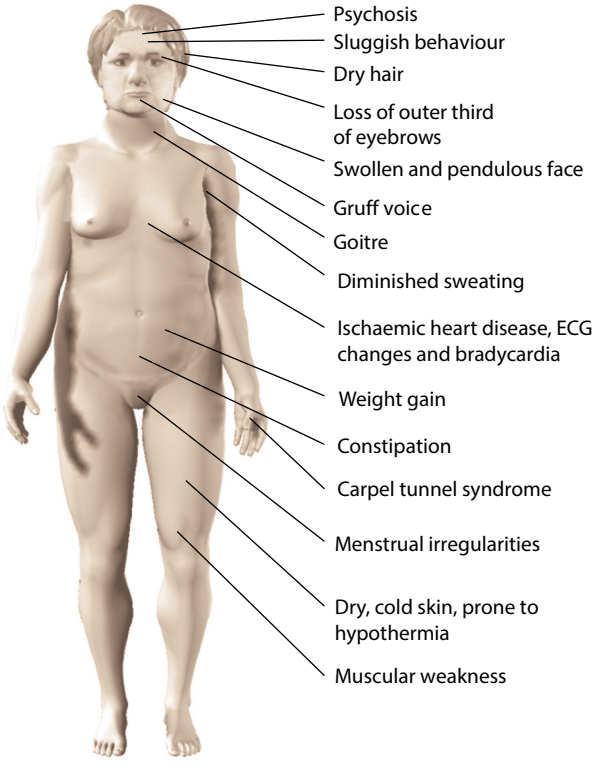


Figure 7.35 Schematic to show the clinical features of myxoedema. ECG, electrocardiogram. ⏎

Diagnosis and treatment of thyroid disorders

Investigating thyroid abnormalities involves measuring serum TSH as a first-line test for thyroid function. Some clinical/pathology laboratories also include measurement of free T_4 and/or free T_3 in their first-line screen. Sensitive assays for measurement of TSH are readily available. Thyroid-stimulating hormone concentrations are increased in primary hypothyroidism, normal in euthyroid individuals (those without thyroid disease) and low or undetectable in hyperthyroidism.

Total T_4 and T_3 measurements were once used widely in pathology laboratories but have the disadvantage that their values are dependent on plasma TBG concentrations, which can give misleading results. When TBG concentrations increase, for example in pregnancy or in females receiving oestrogen-containing oral contraceptives (see *Box 7.5*), then total T_4 and total T_3 concentrations are also increased, even though the individual is not hyperthyroid. Decreases in TBG concentrations occur in malnutrition, protein loss, severe illness and malabsorption (*Chapters 10* and *11*) and cause a reduction

in total T_4 and T_3. Most laboratories now determine free T_4 and T_3, rather than the total concentration. Free T_4 and particularly free T_3 are usually increased in hyperthyroidism. Free T_4 is low in hypothyroidism and is the preferred measurement to aid diagnosis because free T_3 can be normal in 50% of patients with hypothyroidism due to an increase in its peripheral formation from T_4. In a few patients with hyperthyroidism, free T_4 is within the reference range but free T_3 is increased and TSH is nearly always undetectable. This form of hyperthyroidism is referred to as T_3 toxicosis.

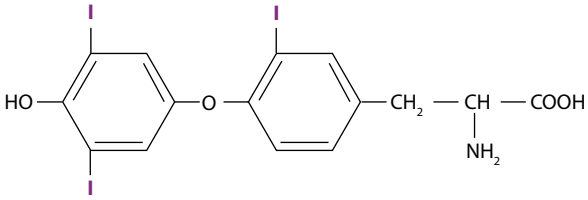


Figure 7.36 Reverse triiodothyronine (rT_3). Compare its structure with those shown in *Figure 7.32*.

In any systemic illness, such as myocardial infarction, fever or liver disease, the normal metabolism of thyroid hormones is disturbed, reducing the concentrations of T_4 and T_3 in the plasma because T_4 is converted to an inactive isomer called reverse T_3 (rT_3; *Figure 7.36*) and T_3 is not replenished from T_4. Thyroid-stimulating hormone concentrations may be normal or reduced and concentrations of TBG, albumin and prealbumin may also decline. Patients may have reduced T_4, T_3 and TSH, although thyroid function is normal, a condition referred to as sick euthyroid syndrome or nonthyroid illness. For this reason, thyroid function tests should not be performed on sick patients until they recover. *Table 7.3* outlines the results of tests used in thyroid disorders. The TRH test is exclusively used in the diagnosis of patients with pituitary disease and to assess the capacity of the pituitary to secrete TSH. The patient is given 200 μg of TRH intravenously and the serum TSH is determined after 0, 20 and 60 minutes. A normal response involves a three- to fivefold increase in TSH above basal levels. A slow rise in TSH (where the 60-minute concentration is greater than that at 20 minutes) together with low basal concentrations of TSH and thyroid hormones suggests hypothalamic disease, while a lack of response is suggestive of pituitary hypothyroidism or hyperthyroidism (*Figure 7.37*). Other techniques for investigating thyroid function include administration of isotopes, such as Tc-99m pertechnetate, and determining their distribution using a camera that detects γ radiation. This technique distinguishes

Test	Hyperthyroidism	Hypothyroidism	Compensated hypothyroidism	T_3 toxicosis	Nonthyroidal illness
TSH	Decreased	Increased	Increased	Undetectable	Decreased or normal
Free T_4	Increased	Decreased	Normal	Normal	Decreased
Free T_3	Considerable increase	Decreased or normal	Decreased or normal	Considerable increase	Decreased

T_3, triiodothyronine; T_4, thyroxine; TSH, thyroid-stimulating hormone.

Table 7.3 Interpretation of results for thyroid function tests.

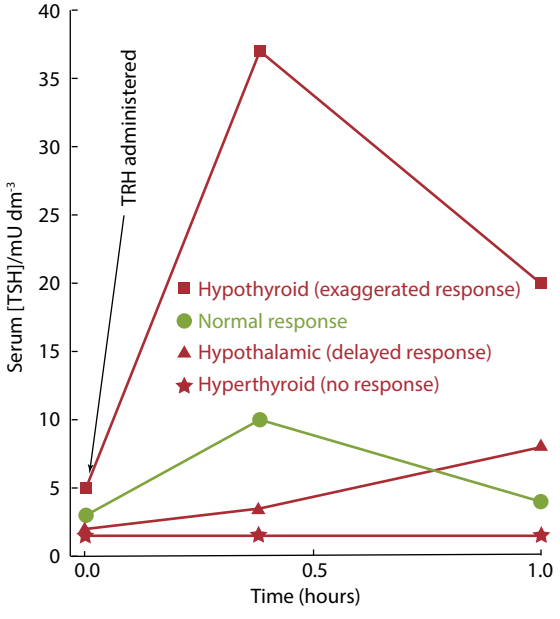


Figure 7.37 Potential results from a thyrotrophin-releasing hormone (TRH) test. TSH, thyroid-stimulating hormone.

between active and inactive thyroid nodules and can distinguish between Graves disease, multinodular goitre or an adenoma affecting the thyroid gland. A thyroid biopsy involves aspirating tissue from the affected region of the thyroid using a syringe and fine needle. The collected cells are examined microscopically for evidence of thyroid nodules. Often, thyroid disease has an autoimmune basis and measurement of antibodies can aid diagnosis. Antiperoxidase and antithyroglobulin antibodies are found in patients affected by Hashimoto thyroiditis and often in those with Graves disease. However, the detection of antibodies is not always diagnostic, as low levels of these antibodies can occur in older people who are euthyroid.

The management of hyperthyroidism includes using antithyroid drugs, for example carbimazole (*Figure 7.38*). This is useful in young patients and acts by reducing the production of thyroid hormones. Other forms of treatment include radioiodine therapy with ^{131}I and partial or complete surgical removal of the thyroid gland (thyroidectomy). Some patients develop hypothyroidism, often because of treatment, and may have to be placed on T_4 therapy. In this regard, it is important to monitor TSH concentrations to detect developing hypothyroidism. Management of hypothyroidism involves replacement therapy with thyroxine, often for life. Thyroxine is readily available, safe and inexpensive. The treatment is monitored at regular intervals by measurement of serum concentrations of TSH to ensure they are kept within its reference range.

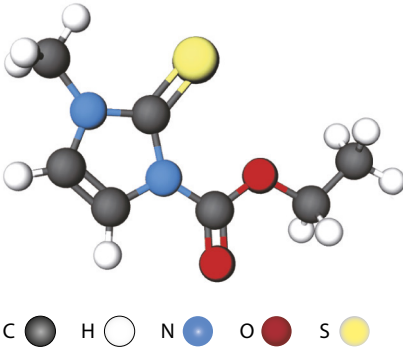


Figure 7.38 Computer drawn model of carbimazole.

7.8 REGULATION OF BLOOD GLUCOSE

Glucose is the universal biological fuel. Indeed, adequate concentrations of glucose in the blood are necessary for brain cells, as they cannot metabolise substances other than glucose and ketone bodies (acetoacetate, acetone and β hydroxybutyrate; see *Figure 9.4*) as energy sources, nor can they store or synthesise glucose. After a meal, any released glucose is absorbed by the gastrointestinal tract (*Chapter 11*), enters the bloodstream and is delivered to the peripheral tissues where it may be metabolised to allow ATP production. Surplus glucose is converted to glycogen and stored in the liver and skeletal muscles or converted to triacylglycerols and stored in adipose tissue. During fasting, the liver produces glucose by glycogenolysis or gluconeogenesis and this is used to maintain blood glucose concentration (*Figure 7.39*).

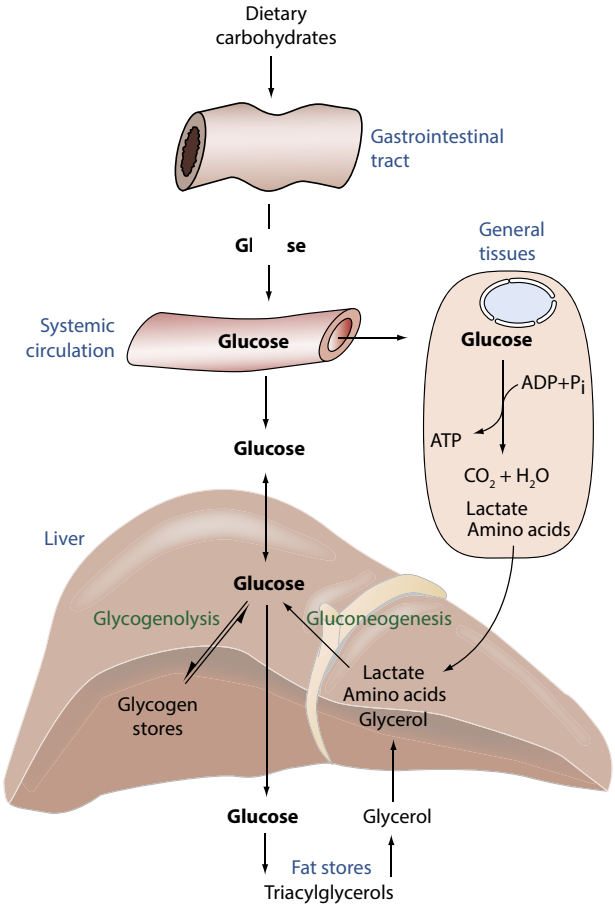


Figure 7.39 Overview of glucose homeostasis.

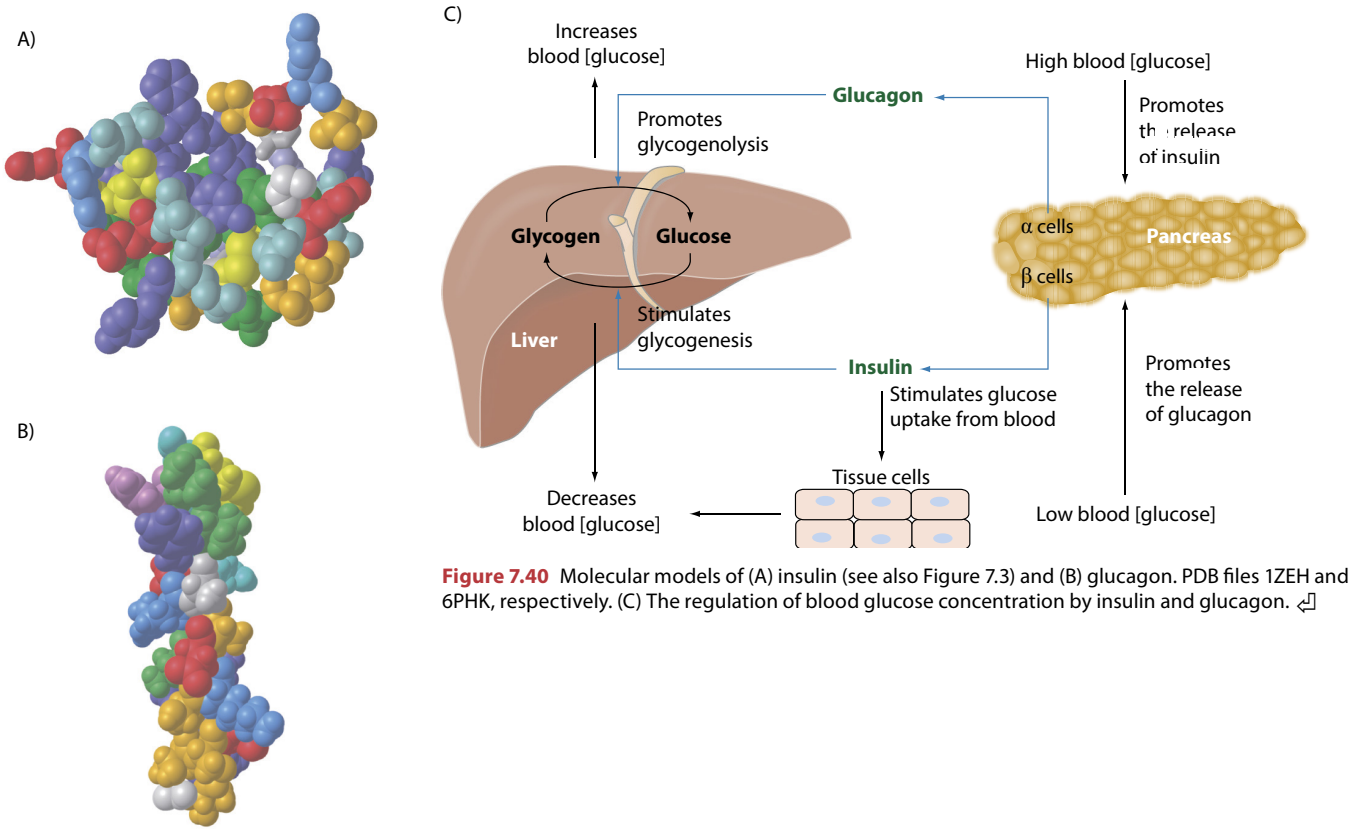


Figure 7.40 Molecular models of (A) insulin (see also Figure 7.3) and (B) glucagon. PDB files 1ZEH and 6PHK, respectively. (C) The regulation of blood glucose concentration by insulin and glucagon.

The concentration of glucose in plasma is regulated by the hormones insulin and glucagon (*Figure 7.40*). Insulin is biosynthesised as the inactive protein preproinsulin in the β cells of the islets of Langerhans of the pancreas, but during its secretion is enzymatically converted to active insulin (*Figures 7.3* and *7.41*). An increase in the concentration of glucose in the blood stimulates the pancreas to secrete insulin. The binding of the hormone to its specific RTK insulin receptor (IR) activates an autophosphorylation as described in *Section 7.3*, which then initiates both rapid and long-term intracellular effects using a cascade of kinases. For example, a major function of the insulin-stimulated signal transduction pathway is to promote the translocation of the insulin-responsive glucose permease GLUT 4 (*Figure 7.42*) from the *trans*-Golgi, endosomes (*Chapter 16*) and GLUT 4 storage vesicles to the plasma membrane, which allows a rapid uptake of glucose by the cells and reduces the concentration of the sugar in blood. More longer-term biochemical and physiological effects are associated with the phosphorylation of transcription factors and the associated changes in gene expression. For example, glycolysis, glycogenesis, protein synthesis and food intake uptake are stimulated. In contrast, glycogenolysis, gluconeogenesis, lipolysis, ketogenesis and proteolysis are inhibited.

Glucagon is released by the α cells of the pancreas. Its effects are antagonistic to those of insulin: reducing concentrations of blood glucose stimulates its release (see *Figure 7.40*). Glucagon binds to a specific GPCR and exerts its effects by activating adenylate cyclase (see *Figure 7.11*) and stimulating a series of protein kinases in the liver; the effect is to promote glycogenolysis and the release of glucose from the cells, which increases the concentration of the sugar in the blood (see *Figure 7.40*).

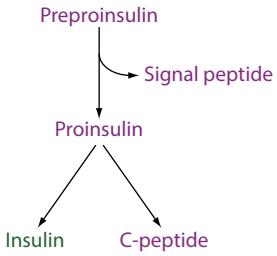


Figure 7.41 The conversion of preproinsulin to active insulin. The signal peptide is the initial amino terminal portion of preproinsulin that allows it to cross the endoplasmic membrane (*Chapter 16*). The C-peptide is an intervening portion of proinsulin that is enzymatically removed, resulting in the formation of an active insulin molecule consisting of two peptides held together by disulphide bonds. See also *Figure 7.3*.

Diabetes mellitus

Diabetes mellitus is a syndrome characterised by hyperglycaemia due to an absolute or relative deficiency of, and/or resistance to, insulin. This is the most common endocrine disorder, affecting about 2% of the world's population. Diabetes can be primary when caused directly by malfunction of one or more of the systems regulating blood glucose concentration or secondary as a result of another disease.

Primary diabetes is divided into types 1 and 2. In type 1 diabetes, also known as insulin-dependent diabetes mellitus (IDDM), there is a decrease or absence of insulin production. It occurs in 15% of all diabetics and typically presents acutely during childhood or adolescence, although it can occur at any age. Patients have marked weight loss, and ketoacidosis (see below) can occur readily.

Type 1 diabetes mellitus is an autoimmune disease (*Chapter 5*) and antibodies that attack β cells of the islets of Langerhans in the pancreas have been demonstrated in over 90% of patients. It also has a strong association with certain histocompatibility antigens such as HLA-DR3, DR4 and certain DQ alleles (*Chapters 4* and *6*). Many cases of type 1 diabetes may develop after a viral infection, such as with Coxsackie B (see *Figure 3.1*), which initiates an autoimmune reaction that destroys the β cells of the pancreas.

Type 2 diabetes is also known as noninsulin-dependent diabetes mellitus (NIDDM), where insulin secretion tends to be normal or even increased. It accounts for about 85% of all cases of diabetes, has a gradual onset and tends to occur in middle-aged and elderly individuals. Patients are less likely to develop ketoacidosis. The aetiology of type 2 diabetes has a strong association with obesity. The disease may arise because an abnormal insulin is produced that is not recognised by the IR or due to a lack of insulin receptors, the presence of defective receptors or by a defective secondary messenger system (*Section 7.3*). Type 2 diabetes has a strong familial incidence. For example, if an identical twin develops type 2 diabetes, there is a strong likelihood that the other twin will also become diabetic.

Secondary diabetes mellitus is uncommon and is a consequence of other disorders that involve excess secretion of hormones antagonistic to insulin, such as cortisol in Cushing syndrome and GH in acromegaly. Damage to the pancreas following, for example, chronic pancreatitis or pancreatic surgery may result in secondary diabetes.

Patients with diabetes mellitus suffer a number of symptoms, including excessive thirst (**polydipsia**), the production of large volumes of urine (**polyuria**), unexplained weight loss, blurred vision, tiredness and an increased susceptibility to infections. Diabetic patients are also susceptible to acute and chronic complications.

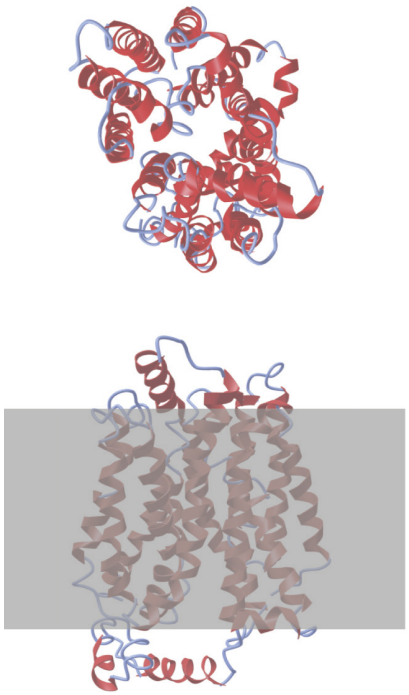


Figure 7.42 Molecular model showing surface and side views of the glucose permease GLUT 4. The approximate position of the membrane is indicated by the grey block. PDB file 4PYP.

Acute complications of diabetes mellitus

Acute complications of diabetes mellitus include diabetic ketoacidosis (DKA), hyperosmolar hyperglycaemic syndrome (HHS, previously known as hyperosmolar non-nonketotic coma [HONK]) and hypoglycaemia. Diabetic ketoacidosis occurs most commonly in patients with uncontrolled type 1 diabetes mellitus as a result of their failure to comply with insulin therapy, but it can be precipitated by infections, such as the common cold, when the body responds by releasing more glucose into the bloodstream and by reducing the action(s) of insulin. The pathogenesis of DKA is outlined in *Figure 7.43*. In DKA, glucose uptake by cells decreases, whereas gluconeogenesis and

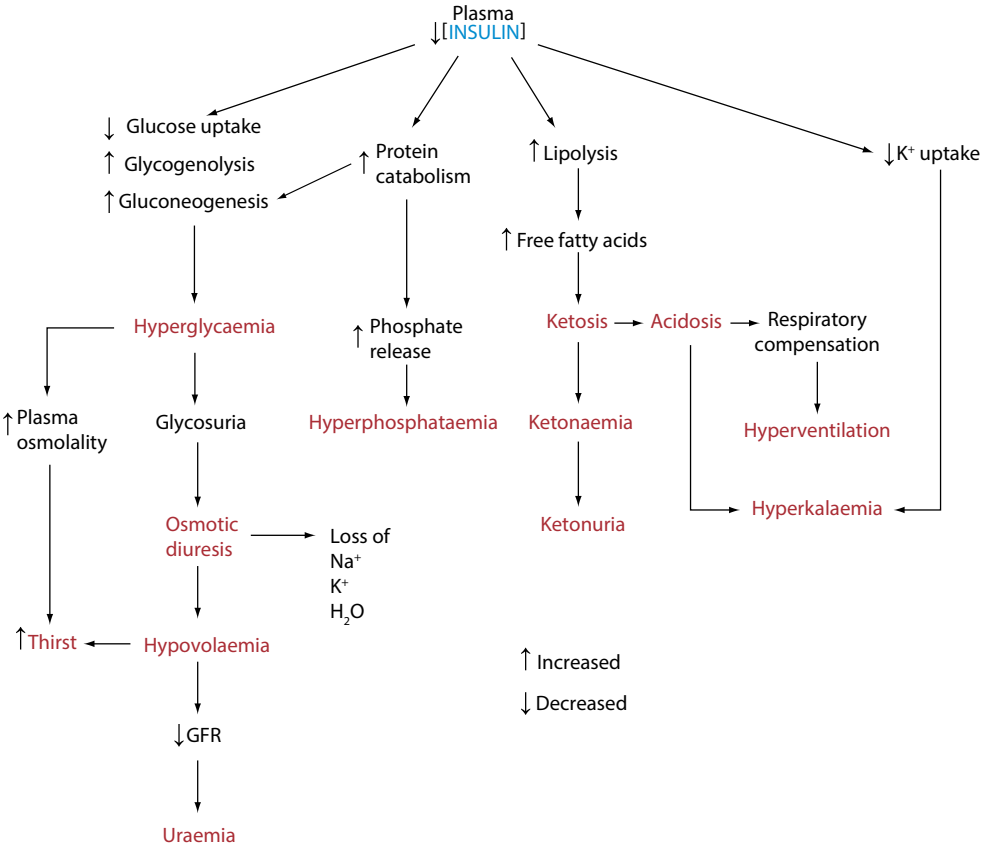


Figure 7.43 An overview of the factors causing diabetic ketoacidosis (DKA). GFR, glomerular filtration rate.

glycogenolysis are both stimulated, causing severe hyperglycaemia. Increased breakdown of proteins occurs, and the released amino acids enter gluconeogenesis or are degraded to form urea. As a result of insulin deficiency, there is a shift of both K^+ and P_i from the intracellular to the extracellular compartments, causing hyperkalaemia and hyperphosphataemia, respectively (*Chapter 8*). The insulin deficiency also stimulates lipolysis producing free fatty acids (FFAs) and glycerol. In the liver, FFAs are converted to acetyl CoA, which is then degraded to the ketone bodies. The liver cannot utilise ketone bodies and they accumulate in the blood (**ketonaemia**) and may be excreted in the urine (**ketonuria**). Ketone bodies in the blood are moderately strong acids, $H^+Ketone^-$, and react with buffers, such as $NaHCO_3$ (*Chapter 9*):

$$H^+Ketone^- + NaHCO_3 \rightarrow Na^+Ketone^- + H_2CO_3$$

decreasing the concentration of hydrogen carbonate but increasing that of carbonic acid in the blood, producing an acidosis. The carbonic acid dissociates to CO_2 and H_2O and the blood PCO_2 increases. Thus, the lung ventilation rates increase as the body attempts to remove the excess CO_2. Ketonaemia is believed to be responsible for the abdominal pain, vomiting and acidosis associated with DKA. Severe hyperglycaemia in DKA exerts a high osmotic pressure, causing water to move out of the cells, leading to cellular dehydration. Blood volume rises and the kidneys respond with polyuria. If blood glucose concentrations exceed the renal threshold, glucose is lost in the urine (**glycosuria**). Eventually, an **osmotic diuresis** may occur with loss of water and electrolytes. The blood volume therefore declines further (**hypovolaemia**) reducing the glomerular filtration rate (GFR) and causing **uraemia**. Dehydration and hypovolaemia may stimulate the thirst centre and the patient may suffer polydipsia.

Diabetic ketoacidosis is a medical emergency and is fatal in about 10% of adults and 5% of children if untreated. Death in untreated cases of DKA is due to poor tissue perfusion, acidosis and cardiovascular failure. The following approaches are taken to correct the metabolic disturbance. Isotonic saline and insulin are administered to correct the dehydration and hyperglycaemia, respectively. Following administration of insulin, K^+ enters the cells and this may cause a hypokalaemia; therefore, it may be necessary to administer K^+ supplements. In the past, severe cases of acidosis were treated by infusion of hydrogen carbonate, although this approach is rarely adopted nowadays. Finally, if the DKA was precipitated by an infection, then it is necessary to identify and treat it.

Patients with uncontrolled type 2 diabetes mellitus may enter an HHS coma. This tends to occur in the elderly and develops over a period of days or even weeks. In these patients, the levels of insulin are sufficient to prevent ketosis, but they often have severe hyperglycaemia, characterised by blood glucose concentration above 35 and often more than 50 mmol/L, together with severe dehydration and high serum osmolality. An HHS coma is usually precipitated by severe illness, diuretics, dehydration and glucocorticoid therapy. Its treatment is similar to that for DKA, except that the patient is rehydrated at a slower rate and insulin requirements are lower than for DKA.

Chronic complications of diabetes

Patients with long-term diabetes, both types 1 and 2, are susceptible to chronic complications afflicting various body organs.

Patients suffering from chronic complications have a reduced life expectancy and, indeed, they are the major cause of morbidity and mortality from diabetes mellitus. Diabetic patients have an increased risk of heart attacks and strokes and, indeed, the former account for most of the deaths in affected patients. Chronic complications also include retinopathy, which is characterised by excessive growth of blood vessels of the eyes and damage to the retina. If untreated, retinopathy can cause blindness. Diabetic cataracts (see *Figure 1.7*) resulting in opacification of the eye lens also result in blindness if untreated. Other complications include diabetic neuropathy with damage to nerves, and this frequently presents as impotence in males. Long-term hyperglycaemia can damage kidneys (diabetic nephropathy), which is characterised by thickening of the basement membrane (see *Figure 8.5*). Impaired wound healing is a serious chronic complication of diabetes: wounds that do not heal properly are susceptible to infections and if not treated in a timely manner can cause gangrene that in serious cases may require amputation.

It is generally accepted that increased protein glycation and the accumulation of tissue advanced glycation end products (AGEs; *Chapter 18*) are involved in the pathogenesis of chronic complications. However, chronic complications are more likely to occur in patients with poor control of their blood glucose concentrations.

Margin Note 7.4
Sanger and sequences

Frederick Sanger (1918–2013) determined the complete sequences of the amino acid residues, the primary structures, of both chains of insulin in the early 1950s. This was the first unequivocal demonstration that proteins have strictly defined primary structures. With the techniques of chemistry and molecular biology available at that time, this was an incredible achievement and Sanger was awarded the 1958 Nobel Prize in Chemistry. Nowadays, sequencing a protein as small as insulin would be a trivial task. Incredibly, a major method for sequencing DNA, the dideoxy method, was also devised by Sanger in 1977, when he and coworkers published the complete sequence of the genome of the virus φX174. Sanger was awarded the Nobel Prize in Chemistry in 1980 for this work. Sanger thus belongs to the tiny elite group of people who have received two Nobel Prizes.

Diagnosis and treatment of diabetes mellitus

Investigations of suspected diabetes involve an examination of clinical features and laboratory tests. A preliminary screening test may identify the presence of urinary glucose, although this is not diagnostic of diabetes mellitus. A patient presenting with symptoms of diabetes mellitus must submit a venous blood specimen and have its glucose concentration determined. A patient is diagnosed as diabetic if the fasting plasma glucose concentration is equal to or greater than 7.0 mmol/L or if the random concentration is greater than 11.1 mmol/L. Diabetes mellitus is excluded if the fasting or random plasma glucose concentrations are less than 6.1 or 7.8 mmol/L, respectively. If the individual under investigation lacks the typical symptoms of diabetes, then a diagnosis cannot be confirmed by a single glucose determination but requires at least one additional positive test on another day or to be investigated using the oral glucose tolerance test (OGTT). The OGTT requires the patient be kept on a normal diet for three days and then fast overnight. A venous blood sample is then taken to determine the basal (fasting) plasma glucose concentration and the patient then drinks 75 g of anhydrous glucose dissolved in a small volume of water. Further blood specimens are collected after one and two hours and their glucose concentrations determined. Plasma glucose values greater than or equal to 7.0 mmol/L for the basal sample or 11.1 mmol/L for the two-hour sample are diagnostic of diabetes mellitus. However, individuals with plasma glucose concentrations less than 7.0 for the basal or 7.8–11.1 mmol/L for the two-hour sample are categorised as having impaired glucose tolerance (IGT). This group has an increased risk of developing cardiovascular disease (*Chapter 14*). Patients with plasma glucose concentrations of 6.1–7.0 mmol/L for the basal sample or less than 7.8 mmol/L for the two-hour sample have impaired fasting glucose (IFG) and are at risk of developing diabetes. Note that the concentrations given relate to venous plasma samples and are different from those for whole blood samples.

The management of diabetes mellitus aims to provide relief from symptoms and reduce the chances of developing acute and chronic complications. This includes educating the patient that diabetes is a life-long disease and affected individuals must be responsible for their own treatment. Regular clinical and laboratory assessment of the patient is required to ensure that treatment is effective, to detect early signs of treatable complications so as to reduce their progression and ensure compliance with treatment. Management involves the dietary restriction of simple sugars and of saturated fats and cholesterol and the ingestion of complex carbohydrates and fibres. Dietary control is

BOX 7.2 DIABETES REMISSION

For many years, type 2 diabetes was regarded as a progressive disease in that its progression deteriorated with time, with increasing β cell failure, reduced insulin sensitivity and more than 50% of patients eventually requiring insulin therapy. Additionally, the risk of developing chronic complications grew as the disease progressed. However, since about 2008, studies have shown that it is possible to reverse type 2 diabetes by placing patients on extremely low energy (calorie) diets. In preliminary studies, diabetic patients were placed on diets of approximately 600 kcal/day (2510 kJ/day) for eight weeks. Consequently, patients lost considerable weight and their body mass index (BMI) and fat mass decreased. More significantly in this context, within a week of the study starting, the concentrations of blood glucose in the participants returned to the nondiabetic range, and within the eight weeks their glycated haemoglobin (HbA_{1c}) concentrations were similarly improved; effectively, the patients had reversed their diabetes. In addition, they showed an improved insulin response following a glucose load and reduced fat contents in their livers and pancreas. Following the eight-week trial, those patients who improved their lifestyle regarding diet and exercise maintained this reversal. Sadly, type 2 diabetes returned in those who reverted to their previous excessive food consumption and sedentary lifestyles.

Initially, *reversing type 2 diabetes* was used to describe this pleasing clinical phenomenon, but the term *diabetes remission* is now preferred because, in some patients in remission, diabetes can return with an increase in blood glucose concentrations. In 2021, international experts from Diabetes UK, the American Diabetes Association and the European Association for the Study of Diabetes formally defined diabetes remission as an HbA_{1c} concentration less than 48 mmol/mol (below 6.5% total haemoglobin) for at least three months without the need for any diabetes medication. Patients with type 2 diabetes achieve remission mainly due to their weight loss. Indeed, obese/overweight (*Chapter 10*) individuals have considerable stores of fats in and around key organs, such as the liver, kidney, heart and pancreas, which can interfere with their functions. Loss of fat improves the activities of these organs, promoting remission. Diabetic patients are more likely to go into remission the sooner they lose weight following their initial diagnosis. However, studies have shown that patients' remission can occur even 10 years after diagnosis. Following a very low energy diet can be challenging, and some clinical monitoring is still required to detect any recurrence of diabetes, particularly in individuals who fail to maintain an appropriate diet and exercise regime. However, diabetes remission offers many advantages. Not only do patients feel better, but any stigma about having diabetes is removed. They have improved blood pressure and cholesterol concentration, and the likelihood of developing chronic complications is reduced, thus improving life expectancy. Less medication is generally required, which reduces the need and costs for the patient and the health service.

often accompanied by the use of injected insulin or oral hypoglycaemic drugs, such as sulphonylureas (*Figure 7.44*), in patients with type 1 and type 2 diabetes, respectively.

Occasionally, it may be necessary to use insulin in patients with type 2 diabetes to control blood glucose effectively. Hypoglycaemic drugs act by increasing the sensitivity of β cells to glucose, therefore stimulating insulin release, or by increasing sensitivity of target cells to insulin. Both effects will reduce blood glucose concentration. Some hypoglycaemic drugs act to reduce the absorption of glucose by the gastrointestinal tract (GIT) or reduce glycogenolysis in the liver.

Diabetic patients on treatment are monitored regularly to ensure that blood glucose is kept in control. Most patients measure their own blood glucose at home regularly using kit methods based on reagent strips and a portable glucose meter (*Figure 7.45*) and adjust insulin dosage according to needs, perhaps following a change in diet, during illness or after exercise. The amount of glycated haemoglobin (see *Figure 1.21*, but see also *Figure 18.7*) in a patient is determined regularly to assess therapy compliance because its presence is an indicator of average glycaemia over the previous six to eight weeks. The amount of glycated haemoglobin (HbA_{1c}) tends to be less than 42 mmol/mol (less than 6% of total haemoglobin) in nondiabetics and equal to/greater than 48 mmol/mol (equal to/greater than 6.5%) in diabetic patients. The target for diabetic patients is a [HbA_{1c}] of 48 mmol/mol (6.5%). Diabetic patients with high values for HbA_{1c} have poor blood glucose control, and their treatment or compliance must be reviewed.

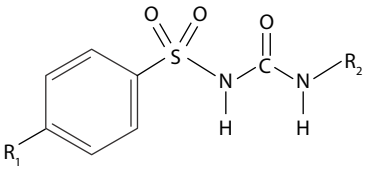


Figure 7.44 Generalised structure of a sulphonylurea.

Figure 7.45 A blood glucose meter and its associated kit.

Hypoglycaemia

Hypoglycaemia is defined as a blood glucose concentration of less than 3.0 mmol/L in a random specimen collected into a tube containing an inhibitor of glycolysis. Hypoglycaemia occurs because of an imbalance between glucose intake, endogenous glucose production and its utilisation. Its clinical features are mainly due to an abnormal functioning of the CNS (neuroglycopenia) and can be acute or chronic.

BOX 7.3 METABOLIC SYNDROME

Metabolic syndrome refers to a combination of obesity, diabetes and hypertension. It is estimated that one in four adults in the UK and 20–25% of those worldwide are affected. The condition places such individuals at a greater risk of having a heart attack or stroke.

Metabolic syndrome is diagnosed if a patient has at least three of the following:

1. they are overweight, particularly with excess fat around the waist (see *Chapter 10*)
2. they have high serum triacylglycerol and high-density lipoprotein concentrations (*Chapter 14*)
3. they have hypertension with a blood pressure of 140/90 mmHg (18.7/12 kPa) or higher (*Chapter 14*)
4. poor control of blood glucose concentration due to insulin resistance

It is possible to prevent or even reverse metabolic syndrome by adopting dietary and lifestyle changes, such as regular exercise, losing weight, stopping smoking, reducing alcohol intake and eating a healthy diet. All these will ensure better control of blood glucose and cholesterol concentrations, and blood pressure.

Acute features include tiredness, confusion, hunger, dizziness, blurred vision, convulsions, coma, anxiety, profuse sweating and tachycardia. The typical signs and symptoms are more likely to occur if blood glucose falls rapidly: in children and young adults, the symptoms often present when the glucose concentration is less than 3.0 mmol/L, whereas neonates develop symptoms only when it is less than 1.5 mmol/L. Acute hypoglycaemia is usually associated with diabetics who have taken too much insulin. Its chronic features include personality changes, memory loss and dementia, and are generally observed in patients with insulin-secreting tumours.

Most cases of hypoglycaemia occur in patients with type 1 diabetes mellitus because of insufficient carbohydrate, too high a dose of insulin, inappropriate use of hypoglycaemic drugs, excessive alcohol intake or strenuous exercise. Hypoglycaemia can occur during fasting if the individual has an insulin-secreting tumour or insulinoma, that is, a primary tumour of the β cells of the islets of Langerhans. These tumours produce excessive amounts of insulin or secrete insulin when it is not required. Nonpancreatic tumours, especially carcinomas of the liver and sarcomas, may cause hypoglycaemia by increasing the cellular uptake of glucose, but this is unlikely to be the sole cause. Many of these tumours secrete IGF-2 (*Figure 7.46*), which has insulin-like effects and is capable of promoting hypoglycaemia.

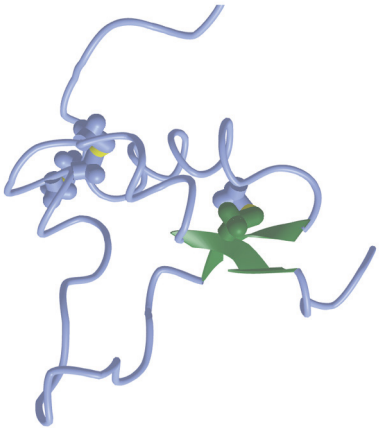


Figure 7.46 Molecular model of insulin-like growth factor-2 (IGF-2). PDB file 1IGL.

Hypoglycaemia in children is particularly dangerous because of the high risk of permanent brain damage. Its diagnosis is especially necessary in the first few months of life. A fetus exposed to maternal hyperglycaemia will have pancreatic islet cell hyperplasia and elevated insulin levels. Following birth, the neonate has hyperinsulinism and may develop hypoglycaemia now the rich glucose supply from the mother has been removed. In addition, these babies are larger than average in size, since insulin promotes growth.

Diagnosis and treatment of hypoglycaemia

An initial assessment of a patient with frequent episodes of hypoglycaemia involves a thorough clinical evaluation with an emphasis on the patient's drug history, relationship of symptoms to meals, presence/history of endocrine disease and an investigation for a possible nonpancreatic tumour. Laboratory tests can confirm the diagnosis of hypoglycaemia by demonstrating a low blood glucose concentration. To confirm that the clinical features are due to hypoglycaemia, the patient can be given glucose by mouth or parenterally as appropriate. Symptoms that are due to acute neuroglycopenia resolve immediately, whereas those due to chronic neuroglycopenia often persist. Measurement of plasma insulin concentration can help in the diagnosis or exclusion of an insulinoma. An insulinoma is likely if the patient has fasting hypoglycaemia together with high serum insulin concentrations greater than 10 mU/L and raised C-peptide concentration (see *Figure 7.41*). Insulin secretion in an insulin-treated diabetic cannot be determined for obvious reasons. However, insulin and the C-peptide are secreted by islet cells in equimolar amounts and therefore a measurement of C-peptide together with insulin can differentiate between hypoglycaemia due to an insulinoma, which will have a high C-peptide concentration, and that due to exogenous insulin, which will have relatively low amounts of C-peptide.

Hypoglycaemia can be fatal, and patients must be treated urgently. The aim in comatose patients is to rapidly correct the hypoglycaemia using intravenous dextrose or intramuscular glucagon. If a patient is conscious and can swallow, then they should be given sweet drinks, sweets or glucose tablets. The underlying cause of hypoglycaemia needs to be identified and rectified; an insulinoma, for example, requires its surgical removal.

7.9 THE ADRENAL CORTEX

Adrenal glands each weighing about 5 g are found on the upper surfaces of the kidneys and consist of an outer cortex and an inner medulla. The cortex consists of three layers: the zona glomerulosa, zona fasciculata and zona reticularis (*Figure 7.47*). The adrenal cortex is essential for life, since it produces three groups of steroid hormones: glucocorticoids, such as cortisol, and adrenal androgens, for example testosterone, are produced by both the zona reticularis and the zona fasciculata, while mineralocorticoids, such as aldosterone, are made in the zona glomerulosa. Adrenal cortex cells have many low-density lipoprotein receptors on their surfaces enabling them to take up precursor cholesterol rapidly from the plasma.

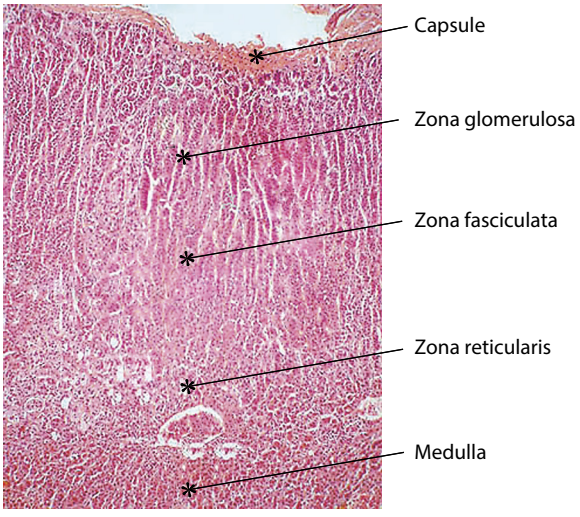


Figure 7.47 The histological structure of the adrenal cortex. Courtesy of Dr A.L. Bell, University of New England, College of Osteopathic Medicine, USA.

Cortisol

Cortisol (*Figure 7.48*) is released in response to ACTH from the anterior pituitary that, in turn, is controlled by the release of corticotrophin-releasing factor (CRF) from the hypothalamus. Cortisol exerts a negative feedback effect on the anterior pituitary and hypothalamus (*Figure 7.49*). The secretion of cortisol shows a diurnal variation, being highest in the morning and lowest at night. Cortisol stimulates an increase in protein catabolism, hepatic glycogenolysis and gluconeogenesis and a redistribution of adipose tissue, but suppresses inflammation. About 90% of blood cortisol is bound to a cortisol-binding globulin called transcortin, while the remaining 10% occurs free.

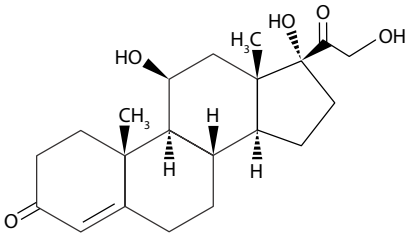


Figure 7.48 Cortisol.

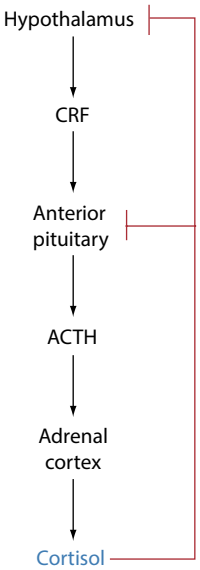


Figure 7.49 The feedback regulation of cortisol secretion. ACTH, adrenocorticotrophic hormone; CRF, corticotrophin-releasing factor.

Aldosterone

Aldosterone (*Figure 7.50*; see also *Figure 8.6*) is the major mineralocorticoid released in response to hypotension, Na^+ depletion or hyperkalaemia (*Chapter 8*). Juxtaglomerular cells in the kidneys detect a fall in blood pressure and secrete renin, which circulates in the blood and catalyses the conversion of the plasma protein angiotensinogen to angiotensin I (*Chapter 14*). An angiotensin-converting enzyme in the lungs converts angiotensin I to angiotensin II, which stimulates the release of aldosterone from the adrenal cortex and ADH from the posterior pituitary (*Figure 7.51*).

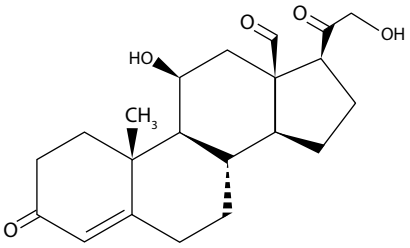


Figure 7.50 Aldosterone. See also *Figure 8.6*.

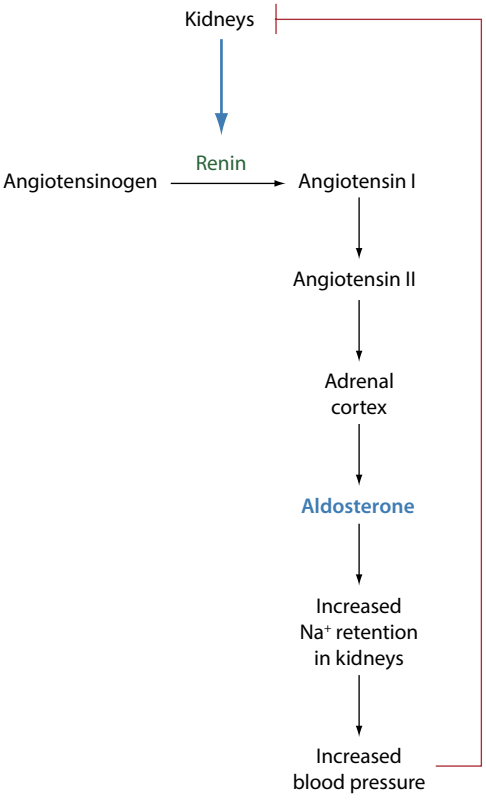


Figure 7.51 The feedback regulation of aldosterone secretion.

Aldosterone stimulates the retention of Na^+ in exchange for K^+ and H^+ in the kidney tubules, increasing the osmolality of the extracellular fluid (ECF) and the retention of water and raising the blood pressure or the ECF volume back to physiological values. Angiotensin II also stimulates thirst and the resultant drinking helps to increase blood pressure. The adrenal cortex secretes androgens, such as testosterone,

dehydroepiandrosterone sulphate (DHEAS) and androstenedione. Testosterone is the hormone that stimulates the growth and development of the male characteristics (*Section 7.10*).

Disorders of the adrenal cortex

Several clinical disorders affect the adrenal cortex. Perhaps the best known are:

- Addison disease
- Cushing syndrome
- congenital adrenal hyperplasia
- Conn syndrome

Addison disease

Addison disease or adrenal hypofunction is a rare condition but is simple to treat once diagnosed. It can arise from one of a number of causes: an autoimmune destruction of the adrenal cortex as a response to tuberculosis (TB), amyloidosis, haemochromatosis, following adrenalectomy, or hypothalamic or pituitary diseases referred to as secondary adrenal insufficiency. Addison disease is characterised by a deficiency of glucocorticoids and mineralocorticoids. Its clinical features include weakness, lethargy, anorexia, nausea, vomiting, weight loss, hypotension, skin pigmentation, hypoglycaemia and depression (*Figure 7.52*). In the first few months, symptoms are usually vague with only lethargy, weakness and weight loss presenting from the glucocorticoid deficiency. Later, patients start to vomit and have abdominal pain. The lack of mineralocorticoids leads to an excessive loss of Na^+ and therefore hypotension is common in these patients. Plasma ACTH is increased because of pituitary response to low levels of cortisol given the lack of a negative feedback mechanism. Adrenocorticotrophic hormone can stimulate melanocytes in skin to produce the melanin; hence, skin pigmentation is common in Addison disease.

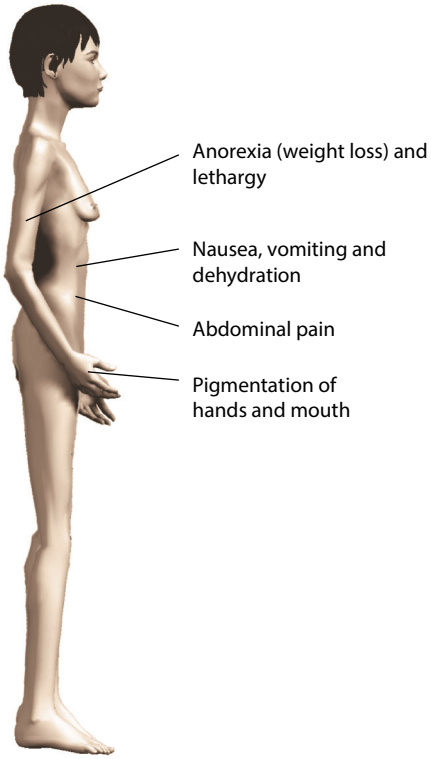


Figure 7.52 Schematic showing the clinical features of Addison disease.

A clinical suspicion of Addison disease can be confirmed by demonstrating hyperkalaemia with hyponatraemia (*Chapter 8*). The plasma cortisol concentration is usually, but not always, low in these patients. A high concentration of ACTH coupled with a low concentration of cortisol is indicative of Addison disease, whereas low cortisol and ACTH values are suggestive of a secondary adrenal insufficiency. The situation can be resolved using complex biochemical tests with the analogue of ACTH, Synacthen (*Figure 7.53*). The short Synacthen test involves injecting 0.25 mg of Synacthen intramuscularly.

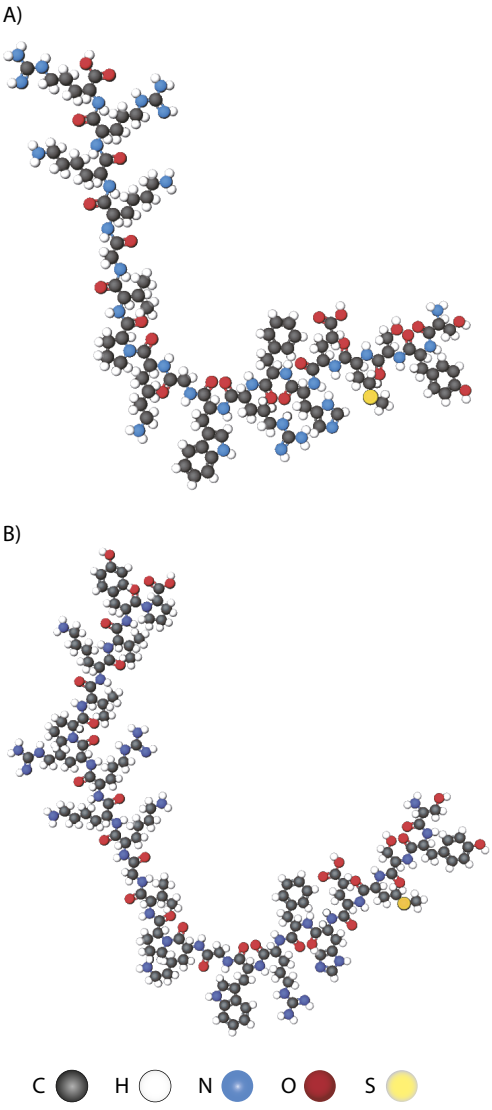


Figure 7.53 Computer drawn structures of (A) adrenocorticotrophic hormone (ACTH) and (B) Synacthen.

The concentration of plasma cortisol is determined within 30 minutes. If it increases by at least 200 nmol/L from the baseline concentration or to a value greater than 550 nmol/L, then Addison disease is unlikely. If this is not the case, then it is appropriate to proceed to the *long* Synacthen test. This involves an intramuscular injection of 1 mg of ACTH daily for three days. On the fourth day, the short Synacthen test is performed and the serum concentration of cortisol is measured. If this is less than 200 nmol/L with no increase following administration of Synacthen, then primary adrenal failure and Addison disease are indicated. If, however, there is an incremental increase of at least 200 nmol/L above the baseline value, then the decreased output of cortisol from the adrenal glands is secondary and due to a deficiency of ACTH caused by a hypothalamic or pituitary disorder. Once Addison disease is diagnosed, it is necessary to determine its cause. A number of clinical laboratories test for the presence of antibodies against the adrenal glands to see if there is an autoimmune cause. A plain abdominal X-ray may be useful in that it can detect calcification of adrenal glands because of TB *(Chapter 3)*.

The conventional therapy for Addison disease involves treatment with steroids, such as hydrocortisone and fludrocortisone (*Figure 7.54*), which possess glucocorticoid and mineralocorticoid activities, respectively. If a patient is left untreated, he or she will eventually experience an adrenal crisis precipitated by stress, bacterial infections, trauma or surgery, which is a medical emergency. Typical clinical features of a crisis are abdominal pain, vomiting, hypotension together with hyponatraemia, hypoglycaemia and hyperkalaemia. Its treatment involves administering saline infusions to correct the hypotension, fluid and salt losses, and intravenous steroids to correct glucocorticoid and mineralocorticoid deficiencies. The precipitating factors, such as bacterial infections (*Chapters 2* and *3*), require identification and appropriate treatment.

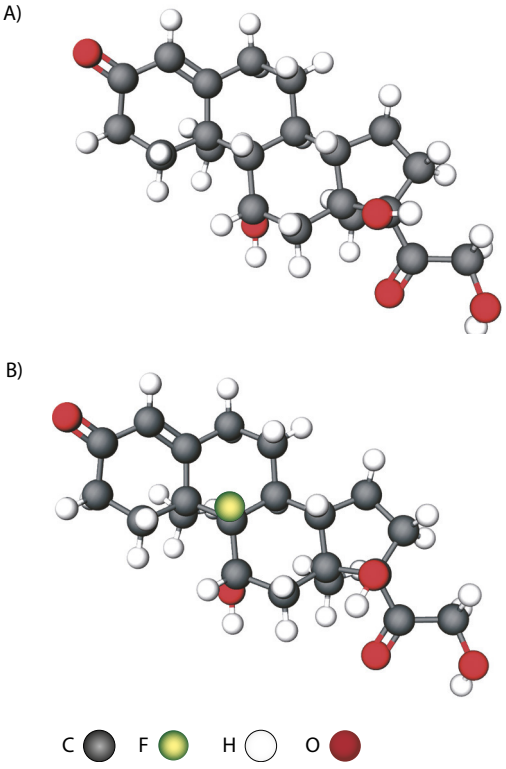


Figure 7.54 (A) Hydrocortisone and (B) fludrocortisone.

Cushing syndrome

Cushing syndrome arises from a number of causes that result in adrenal hyperfunction. The most common cause is a pituitary lesion, which results in the secretion of large amounts of ACTH, a condition referred to as Cushing *disease*. Other causes include ectopic production of ACTH from a carcinoma of the lungs or the excessive production of cortisol from an adrenal adenoma or carcinoma and iatrogenic causes, such as corticosteroid or ACTH treatment. The major clinical features of Cushing syndrome (*Figure 7.55*) include truncal obesity, thinning of skin, excessive bruising, poor wound healing, purple

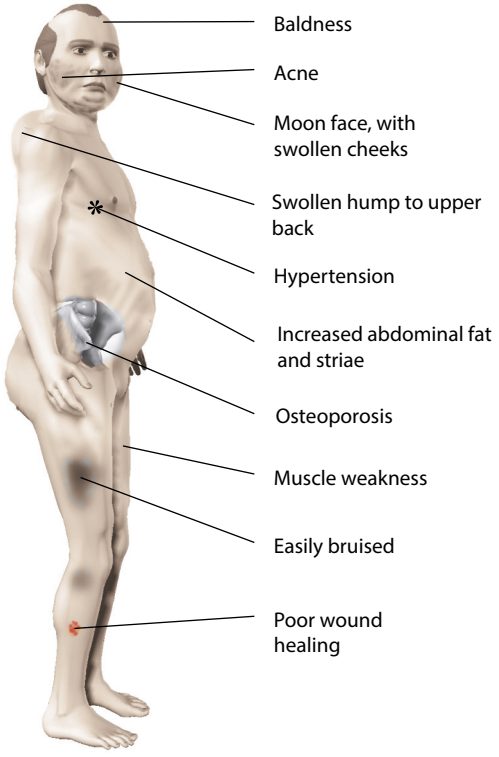


Figure 7.55 Schematic to show the clinical features of Cushing syndrome.

striae on the abdomen and thighs, muscle weakness and wasting, **hirsutism** (the development of increased body hair on the face, chest, upper back and abdomen in females, especially in cases of adrenal carcinoma), hypertension, amenorrhoea and psychiatric disturbances. Excess cortisol has a mineralocorticoid effect leading to the retention of Na^+ and water, producing a hypertension. Hypokalaemia may also occur following an excessive loss of K^+ (*Chapter 8*). Excess cortisol increases blood glucose concentrations and some patients may present with diabetes mellitus. The clinical features are due to increased cortisol production and, partly, to excessive androgen release.

A clinical suspicion of the syndrome is supported by the patient presenting with hypokalaemia and alkalosis, a high urinary free cortisol, which is normally less than 300 nmol daily, and disturbance of the usual diurnal rhythm of cortisol secretion. Initial screening criteria are followed by tests using the cortisol analogue dexamethasone (*Figure 7.56*), which is not detected by the usual methods of measuring cortisol. Dexamethasone, however, suppresses ACTH production and cortisol secretion in healthy people. The low-dose dexamethasone test involves giving 1 mg of dexamethasone at night. A blood specimen is taken for cortisol determination the following morning. A failure of dexamethasone to suppress cortisol release is suggestive of Cushing syndrome or disease. To distinguish between the two, a high-dose dexamethasone test may be used or the concentration of ACTH in plasma measured. The high-dose dexamethasone test consists of administering 2 mg of dexamethasone every 6 hours for a period of 48 hours. The concentration of cortisol in the plasma is then measured at 09:00 in the morning following the last dose. In Cushing syndrome, suppression of cortisol does not occur because of the excessive secretion of cortisol by an adrenal tumour or as a response to an ectopic source of ACTH. In Cushing disease, caused by a pituitary lesion secreting ACTH, the concentration of cortisol is suppressed to less than 50% of its value prior to the test. Plasma ACTH concentrations are increased in patients with Cushing disease and ectopic ACTH production but are low in patients with a cortisol-secreting adrenal tumour.

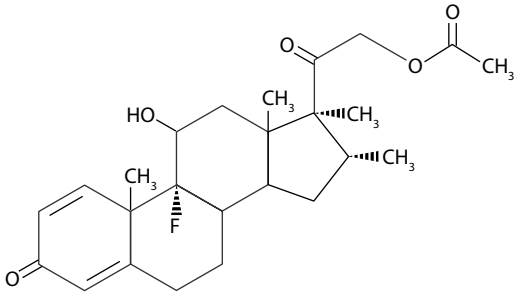


Figure 7.56 Dexamethasone. Compare its structure to that of cortisol (*Figure 7.48*).

The management and treatment of Cushing syndrome depend upon its cause. Drugs, such as metyrapone (*Figure 7.57*), which inhibit the synthesis of cortisol may be used. Tumours can be removed surgically.

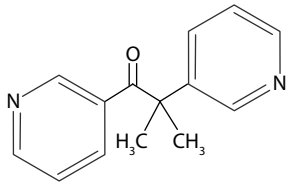


Figure 7.57 Metyrapone.

Congenital adrenal hyperplasia

Congenital adrenal hyperplasia (CAH) is an autosomal recessive condition (*Chapter 15*) characterised by an abnormal biosynthesis of steroid hormones in the adrenal glands. It has an incidence of 1 in 12 000 newborn babies in the UK. Its clinical features depend on whether cortisol and/or aldosterone or androgens are involved. The most common cause of CAH is a complete or partial deficiency of 21-hydroxylase (*Figure 7.58*) that accounts for 95% of all cases. A deficiency of 21-hydroxylase activity blocks the biosynthesis of cortisol (*Figure 7.59*) and, consequently, negative feedback to the anterior pituitary is diminished. The anterior pituitary increases its secretion of ACTH, which causes hyperplasia of the adrenal glands. The substrate 17α-hydroxyprogesterone accumulates and stimulates the production of adrenal androgens. A complete deficiency in 21-hydroxylase activity occurs in approximately one-third of CAH patients. Less common forms of CAH are characterised by deficiencies of other enzymes in the biosynthetic pathway. In patients with a complete deficiency of 21-hydroxylase activity, aldosterone production is also inhibited and these individuals usually present shortly after birth with a life-threatening condition characterised by excessive salt and water loss.

Females affected with CAH are born with ambiguous genitalia, although in cases of partial enzyme deficiency this may not be apparent until early adulthood when the female presents with hirsutism, amenorrhoea and infertility. Males with CAH present

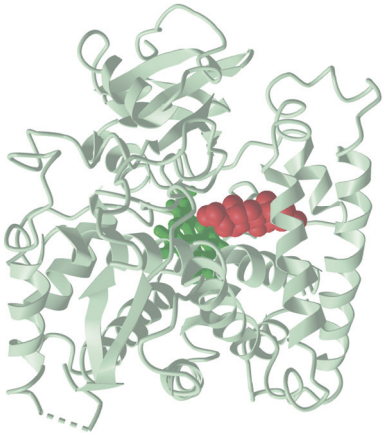


Figure 7.58 Molecular model of 21-hydroxylase. The bound progesterone is shown in red and the haem coenzyme is indicated in the darker green. PDB file 4Y8W.

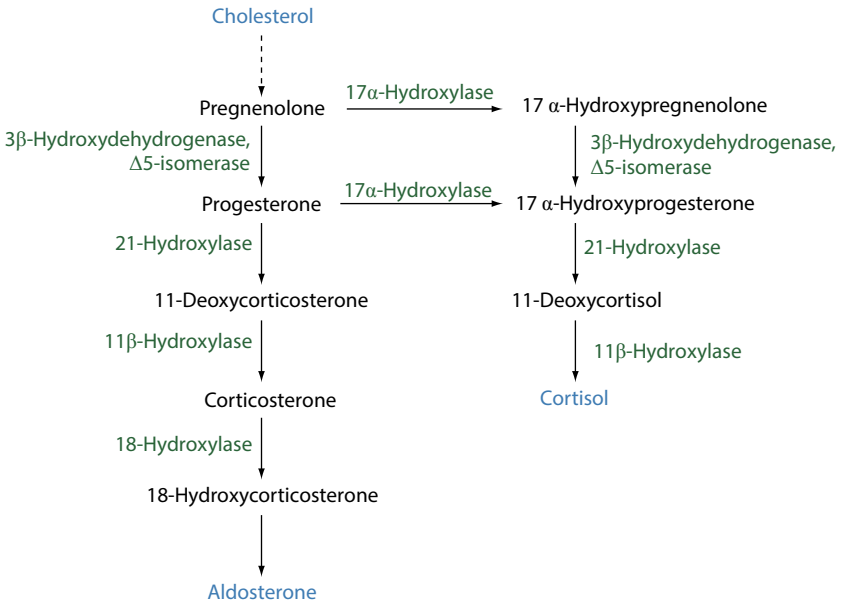


Figure 7.59 The biosynthesis of cortisol and aldosterone. ⏎

with premature development of the male secondary sexual characteristics called pseudoprecocious puberty.

Diagnosis of CAH due to 21-hydroxylase deficiency is made by detecting increased concentrations of 17α-hydroxyprogesterone activity (see *Figure 7.59*) in the baby's blood at least two days following birth. Maternal 17α-hydroxyprogesterone may still be present at two days postparturition, hence the need for the delay. The affected individuals are treated with cortisol and, if necessary, aldosterone. The treatment should reduce ACTH secretion and therefore excessive androgen production. The treatment requires monitoring by regular measurements of plasma 17α-hydroxyprogesterone activity.

Conn syndrome

Conn syndrome, also called primary hyperaldosteronism, is characterised by an increased production of aldosterone. In 80% of cases this is due to an adrenal adenoma. Other causes include hypertrophy of the zona glomerulosa of the adrenal cortex (see *Figure 7.47*) and adrenal carcinomas, although the latter are extremely rare. The excessive secretion of aldosterone leads to the increased retention of Na^+ and loss of K^+ by the kidneys. Most of the clinical features (muscle weakness, tetany, paraesthesiae, polydipsia and polyuria) are due to the hypokalaemia (*Chapter 8*), whereas the excessive Na^+ retention causes hypertension.

Investigating the possibility of Conn syndrome involves determining the concentrations of Na^+ and K^+ in the serum and urine of the patient. Plasma Na^+ can be high, slightly increased or normal, but that of K^+ is always reduced. Measurements of plasma aldosterone and renin in patients with Conn syndrome often show a high aldosterone concentration and a low renin activity. Conn syndrome is managed by treating its underlying cause. Often this is a tumour that must be removed surgically. Other approaches involve using diuretics, such as spironolactone (*Figure 7.60*), an antagonist to aldosterone that helps to control hypertension. This is particularly useful when the cause of Conn syndrome is adrenal hyperplasia.

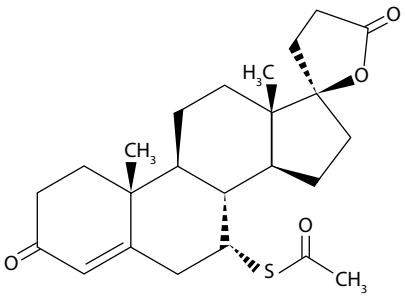


Figure 7.60 Spironolactone. Compare with aldosterone (*Figure 7.50*). ⏎

BOX 7.4 PHEOCHROMOCYTOMA

Pheochromocytomas are tumours of the adrenal medulla that cause uncontrolled and irregular secretion of the catecholamine hormones (*Figure 7.61*). Adrenaline and noradrenaline stimulate increases in heart and respiratory rates and the depth of breathing, dilate blood vessels supplying muscles and increase the metabolism of carbohydrate stores to prepare the body for 'fright, flight or fight'. Pheochromocytomas are usually benign, that is, they do not metastasise (*Chapter 17*), and are rare causes of hypertension (*Chapter 14*), accounting for some 0.5% of such cases.

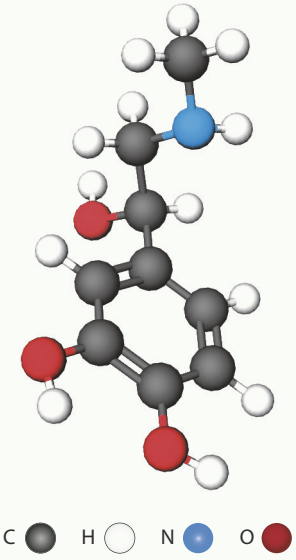


Figure 7.61 Computer drawn structure of adrenaline. See also *Figure 7.3*.

Adrenaline and noradrenaline are metabolised to metadrenaline and normetadrenaline, respectively, by catechol-*O*-methyltransferase (COMT). These products are then converted to 4-hydroxy-3-methoxymandelic acid (HMMA), which is also known as vanillylmandelic acid (VMA), by monoamine oxidase (MAO; *Figure 7.62*). Metadrenaline, normetadrenaline and HMMA are all excreted in the urine.

Patients affected by pheochromocytomas present with headaches, pallor, tremors, increased heart rate and palpitations, sweating and abdominal discomfort. Pheochromocytoma leads to high concentrations of catecholamines in the plasma. However, a diagnosis based on measurement of plasma concentration is difficult, as they are subject to variation by a number of factors including posture. Screening tests include measuring and demonstrating increased concentrations of urinary HMMA or metadrenaline and normetadrenaline (see *Figure 7.62*). The reference range for urine HMMA is less than 35 µmol daily, whereas those for metadrenaline and normetadrenaline are less than 0.1 and below 0.57 µmol/day, respectively. In cases where the results are equivocal, the pentolinium test is conducted. Patients are given 2.5 mg of pentolinium (*Figure 7.63*) by intravenous injection. A blood specimen is collected before and 15 minutes following the injection. Pentolinium inhibits the release of catecholamines in healthy individuals but not in those suffering from pheochromocytoma and thus aids diagnosis.

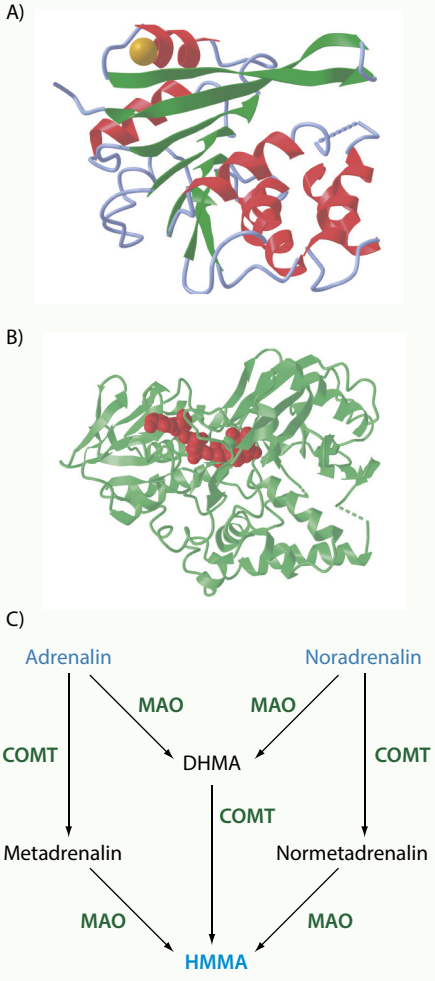


Figure 7.62 Molecular models of (A) catechol-*O*-methyltransferase (COMT; the yellow sphere represents a Na^+) and (B) monoamine oxidase (MAO) with its flavin adenine dinucleotide (FAD) coenzyme is shown in red. PDB files 4PYI and 2BXR, respectively. (C) Overview of the metabolism of adrenaline and noradrenaline.

The treatment of pheochromocytoma involves surgical removal of the tumour. However, it is necessary to block the action of excess adrenaline and noradrenaline before any surgery by administration of phenoxybenzamine (*Figure 7.64*). This is because, during the operation, the sudden release of large amounts of catecholamines would result in high blood pressure that, in turn, could cause a heart attack or brain haemorrhage.

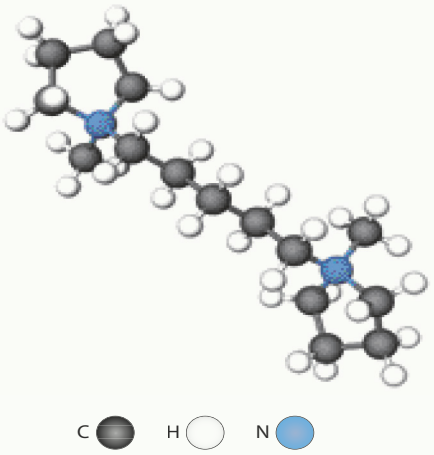


Figure 7.63 Computer drawn model of pentolinium.

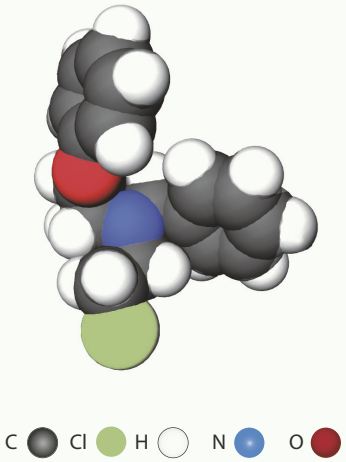


Figure 7.64 Computer drawn model of phenoxybenzamine.

7.10 REPRODUCTIVE HORMONES

Male and female reproductive systems (*Figure 7.65*) produce and secrete several sex hormones that are responsible for the maturation of germ cells, the production of gametes and, in the female, fertilisation of the ovum and its subsequent growth and development.

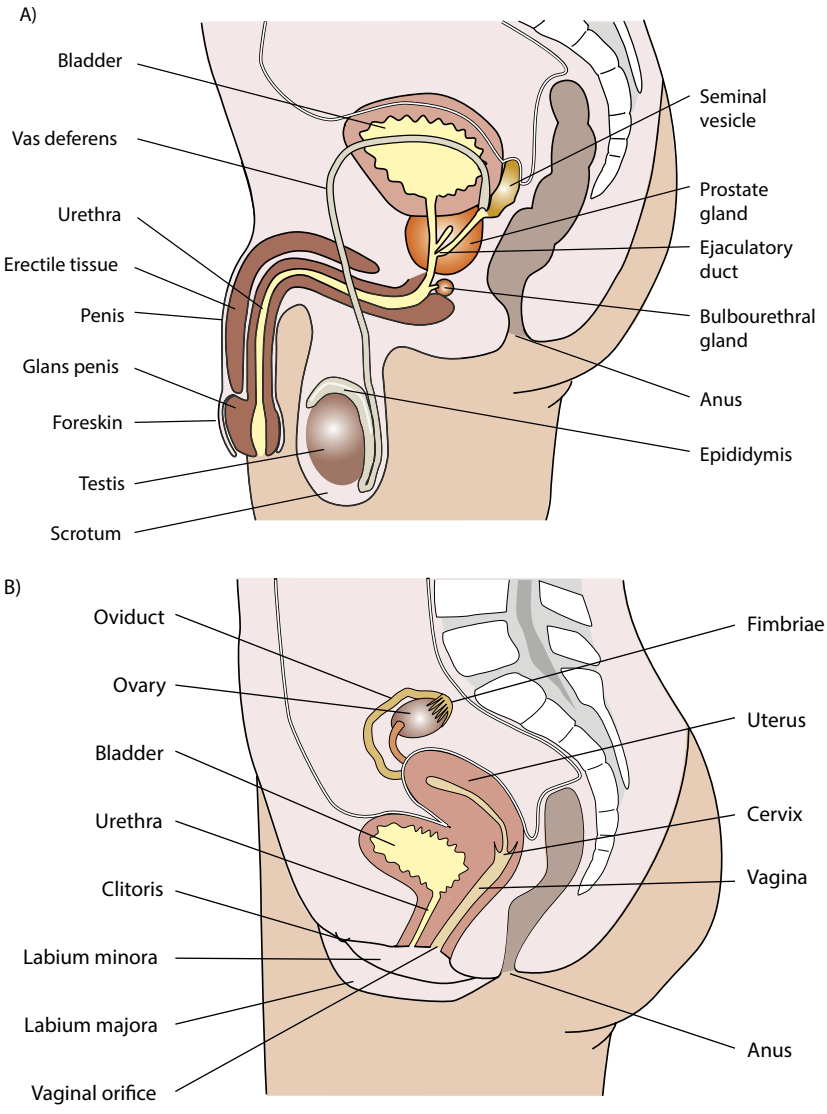


Figure 7.65 Overviews of the (A) male and (B) female reproductive systems.

Testes produce male gametes or spermatozoa ('sperms') that mature and are stored in the epididymis and vas deferens. Testes are composed of lobules with up to three seminiferous tubules containing cells undergoing spermatogenesis. These cells are supported and nourished by Sertoli cells. Spermatogenesis involves meiosis and produces haploid sperm (*Chapter 15*) as outlined in *Figure 7.66*. Each sperm has a head and a tail, the latter consisting of a midpiece and flagellum. The head contains a nucleus and is covered by a cap called an acrosome, which contains enzymes required to penetrate the ovum or egg. The midpiece contains mitochondria that provide energy for the locomotory movements of the flagellum.

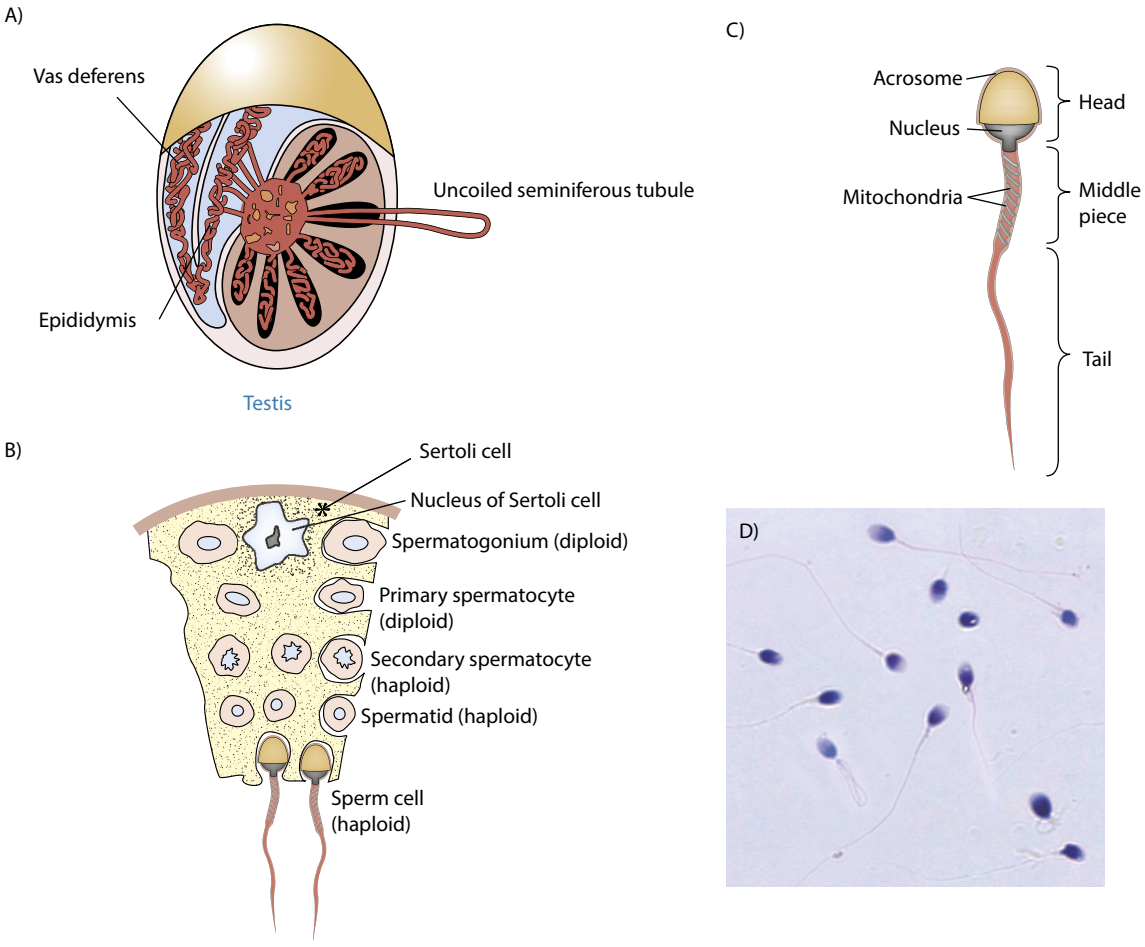


Figure 7.66 (A) Schematic of a testis. (B) Overview of spermatogenesis. (C) Schematic of a spermatozoon (sperm). (D) Photomicrograph of sperm.

The production of ova (oogenesis) or female gametes begins in the ovaries (*Figure 7.67*). Primordial germ cells in the outer germinal epithelium divide by mitosis to form a diploid primary oocyte that becomes surrounded by follicle cells to produce primary follicles. These migrate into the centre of the ovary. As many as two million primary follicles are present at birth and remain dormant until puberty. Approximately 400 primary follicles mature over the lifetime of a woman until follicle development ceases at the menopause. The primary follicle matures to form a secondary follicle. During this development, the primary oocyte divides by meiosis, but this is arrested and forms a haploid secondary oocyte that is the precursor of the ovum and a small polar body. In an adult fertile woman, the nucleus of a secondary oocyte begins the second meiotic division at each monthly ovulation but progresses only to metaphase, when division is again arrested. The second meiotic division of the secondary oocyte is only completed at fertilisation.

In males and females, the hypothalamus secretes GnRH, which regulates secretion of LH and FSH (see *Figure 7.19*) from the basophil cells of the anterior pituitary. Secretion of GnRH, LH and FSH occurs in pulses. Follicle-stimulating hormone and LH act cooperatively to stimulate the ovaries and testes to secrete sex hormones and to develop germ

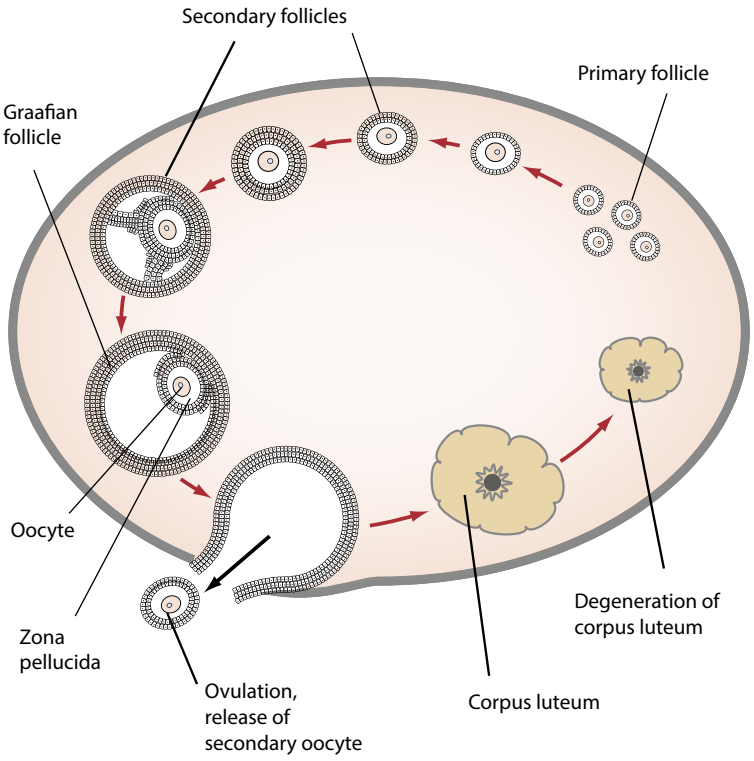


Figure 7.67 Overview of oogenesis and release of oocyte.

cells. Leydig cells in the testes are stimulated by LH to release testosterone, the principal androgen. Its secretion inhibits further release of LH by a negative feedback mechanism. Follicle-stimulating hormone and testosterone are required by Sertoli cells on the basement membrane of the seminiferous tubules to produce inhibin that, in turn, inhibits the secretion of FSH by negative feedback (*Figure 7.68*). Testosterone is also required for sexual differentiation, the development of secondary sexual characteristics and spermatogenesis. It is transported in the blood, bound to sex hormone binding globulin (SHBG; see *Figure 7.13*) and to a lesser extent albumin. Typically, only its free fraction is metabolically active. Testosterone enters target cells and is converted to the potent androgen dihydrotestosterone by the enzyme 5α-reductase. Testosterone is also found

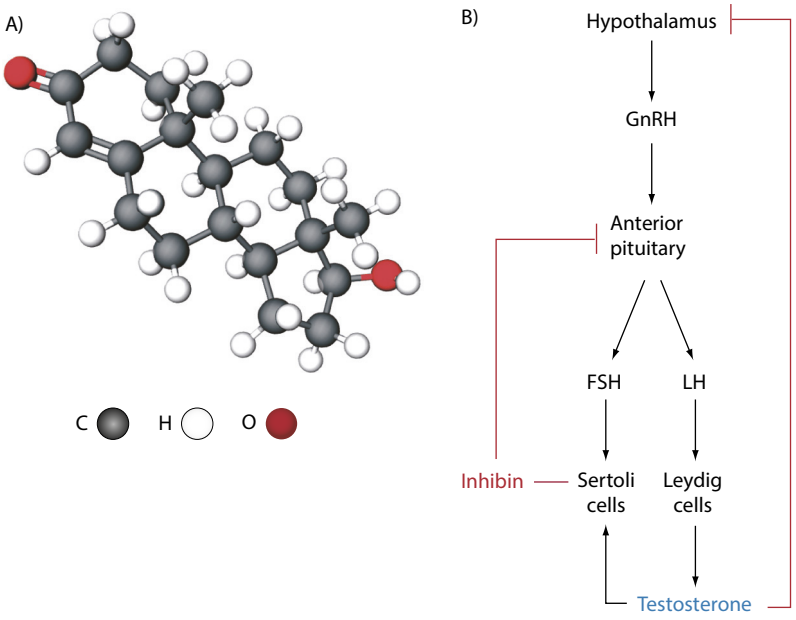


Figure 7.68 (A) Computer drawn structure of testosterone (see also *Figure 7.3*). (B) The regulation of testosterone secretion. FSH, follicle-stimulating hormone; GnRH, gonadotrophin-releasing hormone; LH, luteinising hormone.

in the plasma of normal females, half of which is secreted by the ovaries, the remainder arising from the peripheral conversion of androstenedione and dehyroepiandsrosterone sulphate (DHEAS), both of which are secreted by the adrenal cortex.

The ovaries produce oestrogens, of which oestradiol (*Figure 7.69*) is required for the development of female secondary sexual characteristics and normal menstruation. Circulating oestradiol is bound mostly to SHBG (see *Figure 7.13*), although the blood concentration of oestradiol varies widely with the menstrual cycle. The plasma concentration of oestradiol is low before puberty. It increases rapidly and fluctuates during the menstrual cycle: the series of cyclical changes in the ovary, uterus and pituitary that occur approximately every 28 days until the menopause (see below). Variations in plasma hormones in the menstrual cycle (*Figure 7.70*) depend on interactions between the hypothalamus, anterior pituitary and ovaries.

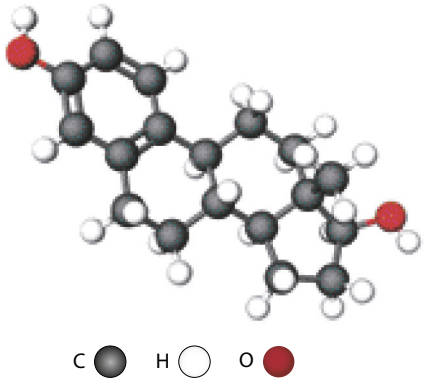


Figure 7.69 Computer drawn model of oestradiol.

Follicle-stimulating hormone is released at the beginning of the cycle and increases growth of the follicles in the ovaries. Oestradiol production increases the sensitivity of the pituitary to GnRH but decreases its secretion by the hypothalamus. The release of oestradiol gradually increases and a follicle matures during the first half of the cycle. At the start of each cycle, about 20 secondary follicles enlarge and begin to secrete oestrogen and the hormone inhibin, and a cavity filled with follicular fluid forms around their ova. This is referred to as antrum formation. By about the sixth day, one of the secondary follicles in an ovary has outgrown the others and becomes the dominant follicle. Its secretion of oestrogen and inhibin decreases the secretion of FSH (see *Figures 7.36* and *7.40*), leading to the regression of the other follicles in an apoptotic process to form atretic follicles. It is uncertain how only the one follicle becomes dominant but appears to be related to its ability to secrete the oestrogen needed for its maturation under the influence of LH. Maturation involves the dominant secondary follicle accumulating fluid-filled cavities that eventually enlarge to the point where they are called a Graafian follicle. Ovulation

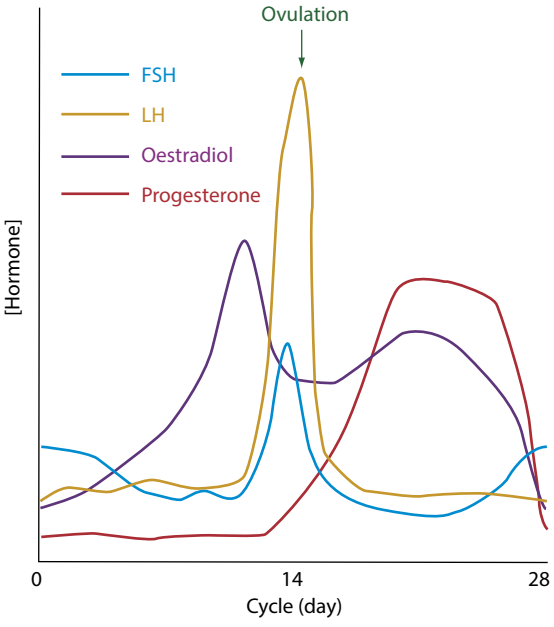


Figure 7.70 Hormonal changes associated with the menstrual cycle. FSH, follicle-stimulating hormone; LH, luteinising hormone.

occurs each month when a Graafian follicle ruptures to release the oocyte, now usually called an ovum, into the fallopian tube. The ovum is transported along the tube by ciliary action. The portion of the follicle remaining in the ovary develops into a corpus luteum. If fertilisation does not occur, this degenerates within 10 days or so.

Menopause

Menopause refers to the stage of life when women stop having periods. It usually occurs between 45 and 55 years of age, with an average age of 51 years, although can happen earlier or later. There is a stage before the menopause called the perimenopause during which the number of mature eggs in the ovaries begins to decline and ovulation becomes increasingly irregular. This is accompanied by reduced production of oestrogen and progesterone. It is the marked decline in the concentrations of oestrogen that causes most of the adverse effects of the menopause in many women. Painful and/or distressing effects include 'brain fog', a colloquial term that rather vaguely describes a mental sluggishness associated with reduced cognitive functioning and mild memory loss when trying to recall, for example, names or dates, hot flushes, difficulty in sleeping, palpitations, worse than usual headaches, muscle and joint pains, dry and itchy skin, vaginal dryness and pain during sexual intercourse. These problems can last for months or years and may even change with time.

A diagnosis of menopause is made on clinical grounds and is based on absence of periods over the past 12 months. Some women choose not to treat their menopausal problems. However, they can be alleviated using hormone replacement therapy (HRT) to give relief from symptoms. Hormone replacement therapy may involve use of combined progesterone and oestrogen (combined HRT) or just oestrogen alone (oestrogen-only HRT). The latter is recommended only if the woman's womb was removed during hysterectomy. Usually, HRT is prescribed in the form of pills, but oestrogen alone can be given as skin patches or vaginal cream.

Fertilisation and pregnancy

Following copulation, the sperm (see *Figure 7.66*) are propelled through the vas deferens by muscular contractions into the urethra. Sperm are suspended in liquid semen produced by the seminal vesicles, prostate and bulbourethral glands. Semen contains nutrients, which activate and increase the motility of sperm, and is alkaline to counteract the acidity of the vagina. The ruptured follicle develops into the corpus luteum, which secretes progesterone and oestradiol and stimulates the development of the endometrium for implantation. Fertilisation of the egg to form a zygote (*Figure 7.71*) usually takes place in the fallopian tubes, and the developing embryo is transported to the uterus by ciliary action and muscular contractions. The zygote begins a series of mitotic divisions to form a developing embryo that embeds into the endometrium lining the uterus and undergoes further development to produce a fetus (*Figure 7.72*), and eventually a neonate in nine months.

Figure 7.71 Photomicrograph of a fertilised ovum (zygote). Courtesy of the Wellcome Collection/Professor A. Handyside.

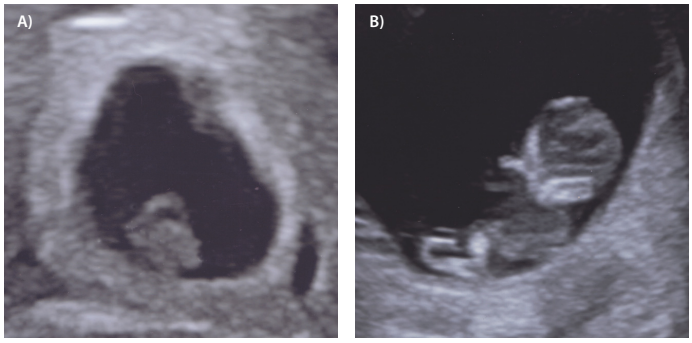


Figure 7.72 Ultrasound scans of a human (A) 7-week-old embryo and (B) 11-week-old fetus. Courtesy of the mother.

Fertilisation ensures that the corpus luteum does not degenerate but begins to produce a number of sex hormones, together with those produced by the gonads and anterior pituitary.

BOX 7.5 ORAL CONTRACEPTIVE PILL

Oral contraceptives (*Figure 7.73*), known colloquially as *the pill*, contain a synthetic version of oestrogen, called oethinyloestradiol, and the synthetic version of progesterone, progestogen. The oestrogen prevents ovulation taking place, whereas progestogen acts on the pituitary gland to block the normal physiological control of the menstrual cycle. Progestogen alters the lining of the uterus so that it is unsuitable for implantation and increases the viscosity of the mucus in the cervix, so that conception is less likely even if ovulation does occur. Oral contraceptive pills are taken daily for three weeks and then stopped for a week during menstruation.

Figure 7.73 Examples of some types of oral contraceptive pills. The inset shows a packet of morning-after pills. Courtesy of the Young Person's Sexual Health Clinic, Brook Advisory Centre, Manchester, UK.

Gregory Pincus (1903–1967) began development of the contraceptive pill in 1950. Within a few years, clinical trials on 6000 women began in Puerto Rico and Haiti. The first commercially available contraceptive pill was introduced in 1960 after the discovery that Mexican yam (*Pachyrhizus erosus*) was a cheap natural source of the hormone precursors required to make the pill. In 2019, over 219 million women worldwide used the pill, with over three million in the UK.

Early contraceptive pills contained between 100 and 175 μg of oestrogen and 100 mg of progestogen. However, shortly after the introduction of the pill, some concern was expressed about side effects. These included an increased disposition to blood clots, heart attacks and strokes, although the risks involved were still relatively small. Studies by 1969 showed that the increased risks of heart attacks and stroke were related to the amounts of oestrogen in the pills. As a consequence, the amounts of oestrogens used have decreased over the years and, by 2006, contraceptive pills contained less than one- third of that in earlier versions. Indeed, a reduction in the risk of heart disease and stroke has been detected in females on modern versions of the pill. Initially, there was some concern that usage of the pill increased the likelihood of cancers of the breast and cervix, but clinical data have cast doubt on this. Using the pill is still associated with some side effects, such as nausea, bleeding between menstrual periods and depression. The pill does increase the chance of blood clot in the legs (deep vein thrombosis (DVT); *Chapter 14*), although the risk for most pill users is very low. It is now known that usage of the pill may have many benefits, for example protection against pelvic infection because the thickened cervical mucus acts as a barrier to bacteria. Also, long-term usage of the pill has been reported to reduce the risk of certain ovarian cancers and can prevent ectopic pregnancies.

The development of the contraceptive pill has been a remarkable achievement, as it allows women to control their fertility in a safe and effective manner. Since its introduction, the pill has had a tremendous impact on female liberty and has aided the process of making pregnancy and motherhood a choice for women. This, particularly in the developed world, has given them greater choices in marriage, work, love and lifestyle.

The 'morning-after pill', also known as the 'postcoital pill', is used by women to reduce the chances of pregnancy following unprotected sexual intercourse. It contains the active ingredient levonorgestrel, which is a synthetic derivative of progesterone. The precise mechanism of action of this pill is still unclear but it is believed to act by preventing ovulation, fertilisation and implantation of the fetus. The whole process from fertilisation to implantation in the womb can take up to three days. Hence, the morning-after pill can prevent pregnancies occurring for up to 72 hours after sexual intercourse, although is more effective the earlier it is taken after copulation. Indeed, it is estimated that 85% of pregnancies would be prevented if the morning-after pill was taken within 72 hours of sexual intercourse.

There is considerable interest in a contraceptive pill designed for use by men. While contraception for women is well established, that for men is limited to using condoms, which have a relatively high failure rate, or vasectomy, which is invasive and cannot easily be reversed. Thus, there is a need for an effective, safe, but reversible, method of contraception for men, similar to the contraceptive pill for women. Studies have shown that testosterone inhibits spermatogenesis, particularly when combined with a progestin, which together suppress the release of gonadotrophin-releasing hormone (GnRH) from the hypothalamus and follicle-stimulating hormone (FSH)/luteinising hormone (LH) from the anterior pituitary. Clinical trials using an injection of these two hormones showed that this approach was successful in lowering the sperm count in men and prevented 96% of pregnancies in their partners. However, the trial was not successful due to unacceptable side effects. However, the future may be more promising. Recent studies in male mice have shown that the compound YCT529 blocks the retinoic acid receptor-α (RAR-α). The binding of retinoic acid to its α receptor is necessary for sperm formation. When binding was inhibited by YCT529, sperm counts were reduced to such low levels that male mice were effectively sterile, and a 99% effectiveness was achieved in preventing pregnancies. No notable side effects were observed. Clinical trials of YCT529 in males are planned; hopefully, this drug will be equally effective in human males and form the basis for an effective (male) contraceptive pill.

Disorders of sex hormones

Disorders associated with male sex hormones include **hypogonadism** and **gynaecomastia**. In the former, there is deficient sperm production and decreased testosterone secretion. Hypogonadism can occur because of testicular disease, referred to as primary or hypergonadotrophic hypogonadism, or a defect in the hypothalamus or pituitary gland leading to secondary or hypogonadotrophic hypogonadism. In the latter, there may be a deficiency of both gonadotrophins or only of LH. The causes of primary hypogonadism are varied. They include congenital defects, such as Klinefelter syndrome (*Chapter 15*), a deficiency of 5α-reductase activity, testicular agenesis (failure of the testes to develop), acquired defects due to testicular infections (for example mumps) or cytotoxic drugs, trauma or irradiation. The causes of secondary hypogonadism include pituitary tumours and hypothalamic disorders, such as Kallmann syndrome (*Margin Note 7.5*). The treatment of hypogonadism is usually directed at the underlying cause. Testosterone is given in cases of testosterone deficiency. However, if fertility is required then gonadotrophins may have to be administered.

Gynaecomastia, or breast development in males, is usually related to a disturbance in the balance of oestrogens to androgens. In puberty, it occurs in approximately 50% of normal boys due to a temporary increase in the secretion of oestrogens. Other than at puberty, the condition is pathological. It may arise because of decreased androgen activity in hypogonadism or because of increased oestrogen production from various endocrine tumours. Such tumours may secrete large quantities of oestrogens or may secrete hCG that stimulates oestrogen production. Some drugs possess oestrogen or antiandrogen activity and their use can lead to gynaecomastia.

Disorders associated with female sex hormones include amenorrhoea, **oligomenorrhoea**, infrequent menstruation and **virilism**, with the development of hirsutism, muscle mass, deepening of the voice and male psychological characteristics. Amenorrhoea can be primary, where menstruation fails to occur by the age of 16 years, or secondary, where menstruation stops for three months or more after normal menstruation has been established and before the menopause. The other clinical features of amenorrhoea include hirsutism, acne, menstrual cycle disturbances and obesity, although these features vary in their severity and prevalence. A common reason for secondary amenorrhoea in females is pregnancy. This condition must be excluded before other possible causes are investigated, for example stress, severe weight loss, polycystic ovary syndrome (PCOS; see below), gonad dysgenesis (such as in Turner syndrome; *Chapter 15*), the decrease in gonadotrophin secretion associated with some tumours, hyperprolactinaemia and congenital adrenal hyperplasia. The causes of amenorrhoea are investigated by measuring the concentrations of FSH, LH and prolactin in plasma (*Figure 7.74*). A high value for FSH

Margin Note 7.5
Kallmann syndrome

Kallmann syndrome (KS) was first described in 1944 by the American psychiatrist Franz Kallmann (1897–1965). Its incidence is estimated at 1 in 48 000 individuals. Kallmann syndrome is a congenital secondary hypogonadism caused by embryological defects in the development of the olfactory mucosa, which results in the defective secretion or, less often, action of gonadotrophin-releasing hormone (GnRH) and decreased levels of sex hormones. In early life, its most notable feature is the absence of the sense of smell (**anosmia**); however, the condition prevents sexual maturation and the development of secondary sexual characteristics and causes infertility. Thus, KS is diagnosed usually when a child fails to begin puberty. Treatment involves lifelong hormone replacement therapy, although male infants may have early hormone treatment or surgery to correct undescended testicles. Unfortunately, later in life, such patients have an increased risk of developing osteoporosis and are often prescribed vitamin D supplements (*Chapter 10*) and bisphosphonates (see *Figure 8.25*).

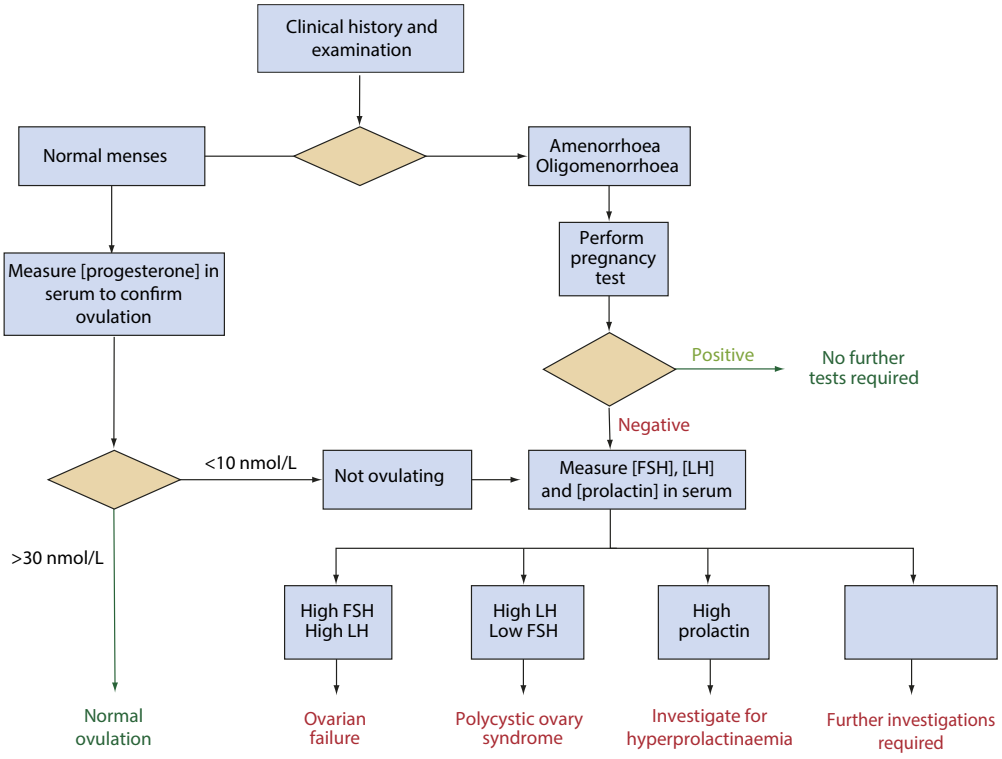


Figure 7.74 Overview of the investigation of amenorrhoea. FSH, follicle-stimulating hormone; LH, luteinising hormone.

indicates ovarian failure. One for prolactin suggests hyperprolactinaemia and requires further investigations to confirm this diagnosis. If, however, the values for FSH, LH and prolactin are normal, further tests to investigate pituitary or hypothalamic diseases are necessary. The management of amenorrhoea is aimed at treating its underlying cause.

Patients with virilism present with enlargement of the clitoris, deepening of the voice, atrophy of the breasts and hirsutism. Hair growth is not only excessive but shows a male-like distribution. The cause of virilism is increased androgen secretion, although abnormally low levels of SHBG can also increase the free testosterone fraction. In some cases, virilism occurs because of an increased sensitivity to androgens by target cells. Its causes include PCOS, androgen-secreting tumours, congenital adrenal hyperplasia and Cushing syndrome and may be iatrogenic following treatment with androgens and progesterone. Its most common cause is PCOS, characterised by multiple cysts in the ovaries that arise from follicles that have failed to ovulate. The ovaries secrete large amounts of androgens, although why this is so is unclear. Many patients with PCOS suffer from acne, obesity and type 2 diabetes mellitus and may be infertile. A diagnosis of PCOS is made on clinical grounds and assisted by ultrasonography. Plasma LH concentration is often increased in these patients but may be normal, whereas that of testosterone is increased. Treatment of severe PCOS may involve using the antiandrogen drug cyproterone, but this leads to infertility. If fertility is required, then clomiphene (*Figure 7.75*), an antioestrogen drug, may be used and can induce ovulation in 75% of cases.

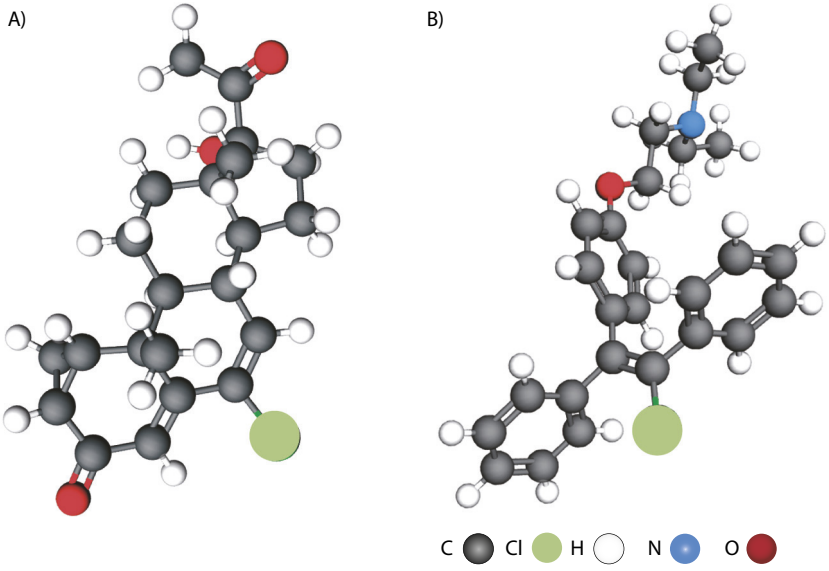


Figure 7.75 Computer drawn models of (A) cyproterone and (B) clomiphene.

Infertility

Infertility is defined as a failure to conceive despite regular unprotected sexual intercourse for one year. Investigating infertility involves a clinical and laboratory evaluation of both partners.

Female infertility may be due to failure to ovulate, obstruction of the fallopian tubes or diseases of the uterine lining. In females, failures to ovulate due to hyperprolactinaemia or to hypothalamic-pituitary dysfunction are responsible for 20% of cases of infertility. Damage to the fallopian tubes may also be a cause of female infertility.

Infertile women should initially be investigated for regular menstruation. If it is regular, then ovulation is probably occurring. The concentration of progesterone in plasma is also indicative and should exceed 30 nmol/L on day 21 of the menstrual cycle; values below 10 nmol/L are suggestive of abnormal ovulation. If ovulation is confirmed then it may be necessary to examine cervical mucus following intercourse to determine the presence of motile, normal shaped sperm.

Male infertility is usually due to the production of sperm with abnormal morphologies, sperm in decreased numbers (oligospermia) or even a complete absence (azoospermia), or sperm with low motility. Infertility due to endocrine dysfunction occurs only rarely in males.

The normal volume of ejaculate is 2–5 mL and contains 40–500 million sperm. A sperm count of less than 20 million per mL is believed to be subnormal and causes 25%

of infertility cases. In general, the lower the total sperm number the greater the chances of infertility, although men with low counts have been known to father children. The motility of sperm is also necessary for fertilisation to take place and at least 60% of sperm should have a normal shape and be mobile with beating flagella. Microscopic examination of samples of ejaculate will indicate whether the sperm produced are motile and normal in shape or show abnormalities, such as distorted tails, malformed heads and twin sperm formed by a failure in development (*Figure 7.76*).

Figure 7.76 Photomicrograph of human spermatozoa (sperm) with various abnormal morphologies, such as bent or coiled tails. Courtesy of Dr M. Carroll, Department of Life Sciences, Manchester Metropolitan University, UK. Compare with *Figure 7.66*.

Sperm counts will show if adequate numbers of sperm are produced. A low sperm count may be further investigated by measuring the concentrations of testosterone, FSH and LH in the plasma to compare with their reference ranges of 9–30 nmol/L, 2–10 U/L and 2–10 U/L, respectively. In some cases, a biopsy of the testes may be necessary. Investigation of infertility in males is outlined in *Figure 7.77*.

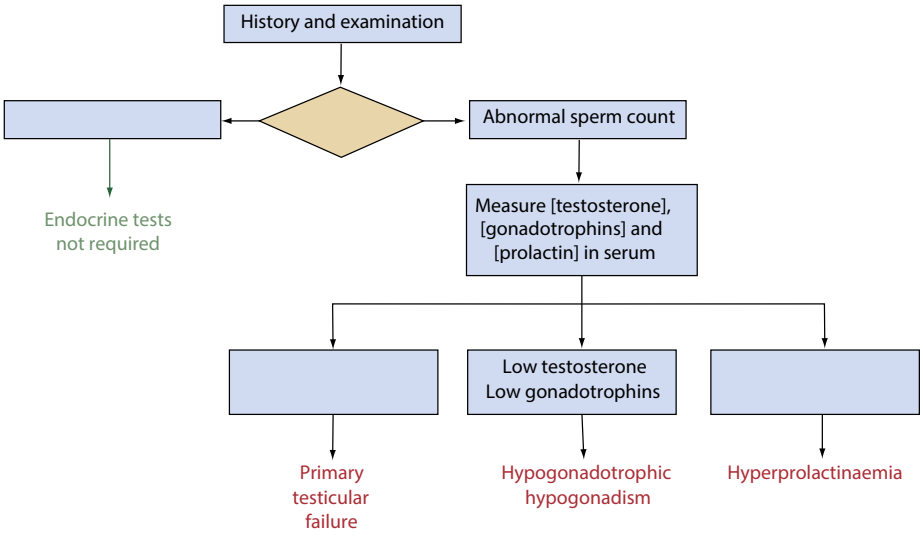


Figure 7.77 Overview of the investigation of infertility in males.

BOX 7.6 ERECTILE DYSFUNCTION AND PHOSPHODIESTERASE TYPE 5 INHIBITORS

Erectile dysfunction (ED) is a persistent inability to achieve or maintain a penile erection necessary for sexual intercourse. The prevalence of ED is difficult to assess. One survey found 7% of men 18–29 years old were affected, which rose to 18% for men of aged 50–59 years. Other studies have shown approximately 50% of men in the 40- to 70-year range had some degree of ED, and almost 10% of these men had complete ED. Indeed, erectile problems may affect up to 4.3 million in the UK and as many as 30 million men in the USA.

Penile erection is governed, at least to a large extent, by the nitric oxide/3',5'-cyclic guanosine monophosphate (NO/cGMP) signalling pathway. Nitric oxide is synthesised by the oxidation of L-arginine in a reaction catalysed by NO synthases (NOs), which requires NADPH and O_2 (*Figure 7.78*). In target tissues, NO enters the cells and binds to and activates soluble guanylyl cyclase (sGC), which catalyses the conversion of GTP to the second messenger, cGMP:

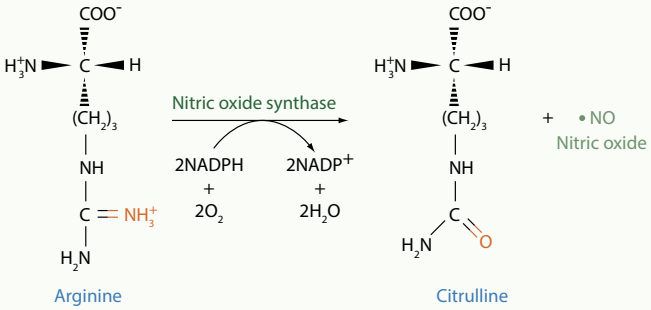


Figure 7.78 The synthesis of nitric oxide (NO) by the catalysed oxidation of L-arginine.

$$GTP + H_2O \xrightarrow{\text{Guanylyl cyclase}} cGMP + PP_i$$

In turn, cGMP activates protein kinase G (PKG), which catalyses the phosphorylation of cation (Ca^{2+}) channels and other proteins leading to decreased intracellular concentrations of Ca^{2+}. In smooth muscle cells, this promotes their relaxation and decreases the tone in blood vessels. This result is reversed when NO and cGMP production ceases and cyclic nucleotide phosphodiesterase (PDE) activity catalyses the hydrolysis of cGMP to GMP:

$$cGMP + H_2O \xrightarrow{\text{Phosphodiesterase}} GMP$$

The arteries and sinusoids that supply the corpus cavernosum and corpus spongiosum ('spongy' tissues collectively called the *corpora*) of the penis with blood contain layers of smooth muscle cells in their walls. In general, this muscle tissue is contracted and restricts the flow of blood into the penis, maintaining it in a flaccid state. However, when suitable sexually aroused, there is an increased release of NO from parasympathetic nerve endings in the walls of the arteries and sinusoids of the corpora, leading to a relaxation of the smooth muscle cells, which allows for an increased flow of blood into the sinusoids increasing their volume. The expanded corpora press against an inflexible layer of collagenous fibres (a tunica albuginea), which reduce the outflow of venous blood from the penis. Hence, the blood pressure in the corpora approaches systolic blood values resulting in penile erection. When erotic stimuli cease, the release of NO from nerve endings declines; the concentrations of cGMP in muscle cells decrease because of reduced biosynthesis and its hydrolysis by the specific phosphodiesterase-5 (PDE5). The smooth muscle cells return to their contracted state and the reduced amount of blood in the corpora allows the penis to return to its flaccid state.

Most cases of ED are believed to be multifactorial, since normal erectile function relies on coordinated endocrine, neurological, muscular and vascular activities, as well as being affected by psychological factors. The concentration of penile cGMP is determined by the relative activities of guanylyl cyclase and PDE5. Hence, the first-line therapy recommended for most men with ED are PDE5 inhibitors (PDEIs), such as sildenafil, although this is better known by its marketed name, Viagra (or colloquially 'the little blue pill'), which has been used clinically since the late 1990s; other PDEIs include vardenafil, tadalafil and avanafil (*Figure 7.79*). The use of PDEIs successfully improves penile erections and sexual intercourse in the majority of sufferers of all ethnic groups. Phosphodiesterase inhibitors are also used clinically to treat other conditions, for example pulmonary hypertension and heart failure (*Chapter 14*). Their adverse effects are relatively mild: diarrhoea, dyspepsia, flushes, headaches, nasal congestion and tinnitus being the most common.

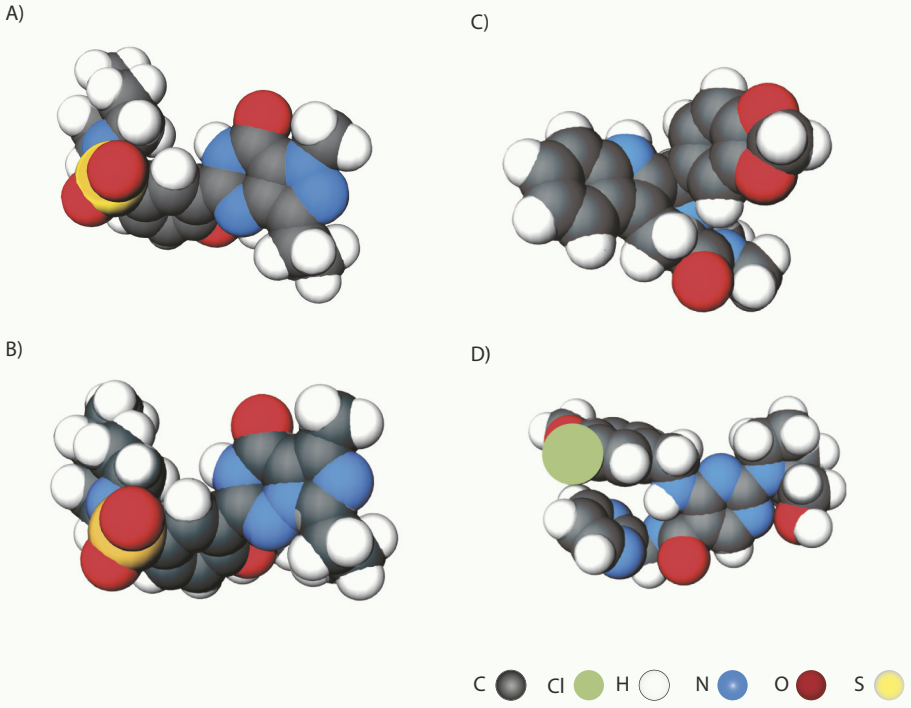


Figure 7.79 Computer drawn structures of the phosphodiesterase-5 inhibitors (PDEIs): (A) sildenafil (Viagra), (B) vardenafil, (C) tadalafil and (D) avanafil.

7.11 CASE STUDIES

CASE STUDY 7.1

Sarah, a 46-year-old woman, was referred to hospital because of her excessive thirst and frequent urination over the preceding two months. Her thirst was severe, and she had to drink water every few hours. An analysis of her plasma and urine yielded the following results (reference ranges are given in parentheses):

Plasma:	Na^+	139	(136–146 mmol/L)
	K^+	4.2	(3.7–5.1 mmol/L)
	Creatinine	105	(65–121 μmol/L)
	Osmolality	296	(282–296 mmol/kg)
Urine:	Glucose		Negative
	Osmolality	90	(up to 1400 mmol/kg)

A fluid deprivation test was performed and the following results were obtained for the plasma and urine osmolality:

Plasma:	Osmolality	310	(282–296 mmol/kg)
Urine:	Osmolality	225	(up to 1400 mmol/kg)

She was allowed to drink water on completion of the test and given a dose of desmopressin after which her urine osmolality increased to 620 mmol/kg.

Questions

(a) What could be the causes of Sarah's symptoms?
(b) Account for the plasma and urine data.
(c) Discuss the need for and results of the fluid deprivation test.

CASE STUDY 7.2

Ian, a 53-year-old man, went to see his family doctor complaining of weakness. Thyroid function tests were requested to exclude the possibility of hypothyroidism. These tests showed the following results (reference ranges are given in parentheses):

TSH	7.5	(0.2–4.0 mU/L)
Free T_4	14.5	(12–25 pmol/L)

Questions

(a) What is the most appropriate clinical explanation of these results?
(b) What other signs and symptoms should be looked for?
(c) How should Ian's progress be monitored?

CASE STUDY 7.3

Amelia, a 20-year-old student, had not been feeling well for three weeks. She was admitted to her local hospital with a history of weakness and vomiting over the previous three days. On admission she was unconscious. Clinical tests on serum yielded the following data (reference ranges are given in parentheses):

Na^+	113	(137–144 mmol/L)
K^+	5.8	(3.3–4.2 mmol/L)
Urea	30	(2.6–6.5 mmol/L)
pH	7.3	(7.35–7.45)

Questions

(a) What is the most likely explanation of these results?
(b) What is the probable cause of Amelia's condition?
(c) Having identified her condition, how should Amelia be treated?

CASE STUDY 7.4

An 18-year-old girl, Jaclyn, was admitted to hospital in a coma. She presented with a deep sighing respiration and her breath smelt of acetone. She had vomited earlier. On presentation, she had a low blood pressure and a high pulse rate of 120 per minute. The following results were obtained (reference ranges are given in parentheses):

Plasma:		
Na^+	136	(132–144 mmol/L)
K^+	5.7	(3.2–4.8 mmol/L)
Urea	15	(3.0–8.0 mmol/L)
Glucose	31.0	(3.0–5.5 mmol/L)
Osmolality	371	(282–295 mmol/L)
Arterial blood:		
pH	7.09	(7.35–7.45)
PCO_2	2.7	(4.7–6.0 kPa)
HCO_3^-	11	(24–34 mmol/L)
Urine:		
Ketones	Positive	Negative

Questions

(a) Explain these results.
(b) How should Jaclyn be treated?

CASE STUDY 7.5

Rachel, a 32-year-old teacher, was suffering from irregular periods and acne. She was examined by her doctor and found to be overweight and hirsute. She was referred for hospital tests, which yielded the following data (reference ranges are given in parentheses):

Plasma:		
Testosterone	3.9	(1.1–3.3 nmol/L)
LH	14	(2.0–10 U/L)
FSH	5.3	(2.0–8.0 U/L)

Questions

(a) What can be inferred from Rachel's symptoms?
(b) What is the most likely diagnosis?

CASE STUDY 7.6

Abdullah, a 58 year old, went to see his general practitioner, as he had experienced headaches over the last six months. On examination, he had coarse facial features with a prominent jaw. His blood pressure was 146/102 mmHg (19.5/13.6 kPa). Serum biochemical investigation gave the following results (reference ranges in parentheses):

Glucose	8.6	(3.5–6.5 mmol/L)
FSH	5.2	(2.0–8.0 U/L)
LH	2.5	(2.0–10.0 U/L)
TSH	2.0	(0.2–4.0 mU/L)
Free T_4	12.9	(12–25 pmol/L)

Abdullah had an oral glucose tolerance test (OGTT) performed, which gave the results shown in *Table 7.4*.

Time (min)	[Glucose] (mmol/L)	[GH] (mU/L)
0	6.9	8
30	13.1	13
60	14.2	13
90	13.8	12
120	11.0	10

Table 7.4 The results of Abdullah's oral glucose tolerance test.

Questions

(a) Explain these results.
(b) How should Abdullah's clinical condition be managed?

7.12 SUMMARY

The endocrine glands and the hormones they produce control many of the activities of the body. Hormones are released into the blood and travel to their target sites of action where they bind to cell surface or intracellular receptors to initiate intracellular activities. Hormones belong to several categories of molecules including amines, peptides, proteins and steroids. Endocrine disorders may arise from damage to the endocrine gland, causing hypo- or hypersecretion of the hormone, or some failure of the hormone–receptor interactions. Disorders of the pituitary gland may arise from pituitary tumours, such as prolactinomas, with excessive secretion of hormones. Abnormal GH secretion leads to dwarfism or acromegaly. Several thyroid disorders arise as a consequence of autoimmune conditions, which can cause hypo- or hyperactivities. However, the most common endocrine disorders involve insulin resistance or its absence, leading to dysregulation of blood glucose concentrations. Disorders of the adrenal gland can result in Addison disease or Cushing syndrome. Dysfunctions associated with reproductive hormones can cause several clinical conditions, including gynaecomastia in men and disruption of menstruation, infertility and virilism in women.

The diagnosis of endocrine disorders usually involves measurement of the hormone in question and replacement for deficiencies. Disorders due to hypersecretion from endocrine tumours may be treated by surgical excision of the tumour.

QUESTIONS

1. Pituitary tumours can secrete various hormones. Which of the following hormones is the most commonly secreted?
(a) TSH.
(b) FSH.
(c) LH.
(d) Prolactin.
(e) GH.

2. Which of the following may cause hyperthyroidism?
(a) Graves disease.
(b) Hashimoto thyroiditis.
(c) Thyroidectomy.
(d) Antithyroid drugs.
(e) Diabetes insipidus.

3. Which of the following can cause acromegaly?
(a) Diabetes mellitus.
(b) Pituitary adenoma.
(c) Excessive ingestion of growth hormone.
(d) Bronchial carcinoma.
(e) Excessive secretion of prolactin.

4. Which one of the following is **NOT** a cause of Cushing syndrome?
(a) Ectopic ACTH production.
(b) Adrenal adenoma.
(c) Corticosteroid therapy.
(d) Autoimmune destruction of the thyroid gland.
(e) All of the above.

5. Ketone bodies are produced during diabetic ketoacidosis as a result of an increase in which of the following?
(a) Protein breakdown.
(b) Urea production.
(c) Insulin release.
(d) Lipolysis.
(e) K^+.

6. Oestradiol is required for which of the following?
(a) The development of male secondary sexual characteristics.
(b) Spermatogenesis.
(c) The development of female secondary sexual characteristics.
(d) The secretion of breast milk.
(e) Stimulating the endometrium to receive a fertilised egg.

7. Sajida, a 27 year old, found over a two-month period that she was expressing breast milk. She was not pregnant and, indeed, had not had any children. She was on medication with analgesic and antidopaminergic drugs for migraine attacks. An analysis of her serum gave a prolactin value of 1200 mU/L (reference range: 70–395 mU/L).
   (a) What is the most likely diagnosis?
   (b) What could have caused such a condition?
   (c) What should her doctor recommend?

8. Graham, a 50-year-old man, visited his doctor for the first time in many years complaining of blurred vision and frequent headaches. The doctor noticed changes in the facial features characteristic of acromegaly. An OGTT including measurements of GH was performed (*Table 7.5*).

Time (min)	[Glucose] (mmol/L)	[GH] (mU/L)
0	6.3	47
30	7.3	55
60	8.6	52
90	9.5	64
120	8.8	53
150	7.5	49
180	6.8	49

Reference range for fasting glucose: 3.0–5.5 mmol/L.
Reference range for GH following glucose load: less than 2 mU/L.

Table 7.5 The results of Graham's oral glucose tolerance test. ⏎

   (a) Plot a graph to illustrate these data.
   (b) Do these findings support the diagnosis of acromegaly?

9. Tabulate the differences between diabetes mellitus types 1 and 2.

10. Give four functions of human GH.

11. Why is it necessary to screen newborn babies to detect congenital hypothyroidism?

12. Which of the following statements is/are **TRUE**?
   (a) Cortisol, aldosterone and testosterone are all secreted by the adrenal cortex.
   (b) With the exceptions of retinoic acid and 1α,25-dihydroxycholecalciferol, all steroid hormones are derivatives of cholesterol.
   (c) Hormones may be released in response to nervous, hormonal or metabolic stimuli.
   (d) Receptors for cytokines and growth factors are generally monomers that dimerise on binding of the cytokine/growth factor.
   (e) Steroid hormones consist of hydrophobic molecules and are transported in the blood bound to specific carrier proteins.
   (f) In the absence of hormone, receptors for steroid hormones are found in the cytoplasm bound to heat shock proteins.
   (g) A possible prolactinoma can be investigated by assessing the concentration of prolactin in the plasma following stimulation with GnRH.
   (h) Growth hormone and somatotrophin are peptide hormones.
   (i) Addison disease is an example of a clinical condition associated with adrenal hyperfunction.
   (j) Nitric oxide is biosynthesised by the oxidation of L-arginine in a reaction that depends on $NADPH^+$ and dioxygen.

FURTHER READING

Aguiar-Oliveira MH and Bartke A (2019) Growth hormone deficiency: health and longevity. *Endocr Rev* **40**:v575–601.

Auer MK *et al*. (2023) Congenital adrenal hyperplasia. *Lancet* **401**:v227–244.

Barthel A *et al*. (2019) An update on Addison's disease. *Exp Clin Endocrinol Diabetes* **127**:165–175.

Calimag APP *et al*. (2023) Diabetic ketoacidosis. *Dis Mon* **69**:101418.

Carson SA and Kallen AN (2021) Diagnosis and management of infertility: a review. *JAMA* **326**:65–76.

Chaker L *et al*. (2017) Hypothyroidism. *Lancet* **390**:1550–1562.

Christ-Crain M and Gaisl O (2021) Diabetes insipidus. *Presse Med* **50**:104093.

Dayan CM *et al*. (2019) Changing the landscape for type 1 diabetes: the first step to prevention. *Lancet* **394**:1286–1296.

ElHady AK *et al*. (2023) Advancements in phosphodiesterase 5 inhibitors: unveiling present and future perspectives. *Pharmaceuticals (Basel)* **16**:1266. doi: 10.3390/ph16091266.

Fahed G *et al*. (2022) Metabolic syndrome: updates on pathophysiology and management in 2021. *Int J Mol Sci* **23**:786.

Farrugia F-A and Charalampopoulos A (2019) Pheochromocytoma. *Endocr Regul* **53**:191–212.

Fleseriu M *et al*. (2022) Acromegaly: pathogenesis, diagnosis, and management. *Lancet Diabetes Endocrinol* **10**:804–826.

Hancock, ML *et al*. (2019) Insulin receptor associates with promoters genome-wide and regulates gene expression. *Cell*, **177**; 722-736e22.

Hughes K and Eastman C (2021) Thyroid disease: long-term management of hyperthyroidism and hypothyroidism. *Aust J Gen Pract* **50**:36–42.

Lee SY and Pearce EN (2023) Hyperthyroidism: a review. *JAMA* **330**:1472–1483.

Majety P *et al*. (2023) Pharmacological approaches to the prevention of type 2 diabetes mellitus. *Front Endocrinol (Lausanne)* **14**:1118848.

Matheson E and Bain J (2019) Hirsutism in women. *Am Fam Physician* **100**:168–175.

Muneer M (2021) Hypoglycaemia. *Adv Exp Med Biol* **1307**:43–69.

Reincke M and Fleseriu M (2023) Cushing syndrome: a review. *JAMA* **330**:170–181.

Shrivastava S and Conigliaro RL (2023) Polycystic ovarian syndrome. *Med Clin North Am* **107**:227–234.

Taylor R *et al*. (2021) Nutritional basis of type 2 diabetes remission. *BMJ* **374**:n1449.

Teal S and Edelman A (2021) Contraception selection, effectiveness, and adverse effects: a review. *JAMA* **326**:2507–2518.

Thiel GT *et al*. (2021) Insulin-responsive transcription factor. *Biomolecules* **11**:1886. doi: 10.3390/biom11121886.

Thirumalai A and Amory JK (2021) Emerging approaches to male contraception. *Fertil Steril* **115**:1369–1376.

Wasilewski T *et al*. (2020) Biochemistry of infertility. *Clin Chim Acta* **508**:185–190.

World Health Organization (2021) WHO laboratory manual for the examination and processing of human semen, 6th edition. WHO Library Cataloging-in-Publication Data.

Yuan Y *et al*. (2022) Cryo-EM structure of human glucose transporter GLUT4. *Nat Commun* **13**:2671. doi: 10.1038/s41467-022-30235-5.

Disorders of water, electrolytes and urate balances

8

OBJECTIVES

After studying this chapter you should be able to:

- describe the structure of the kidney
- list the functions of the kidney
- describe renal function tests
- describe some types of kidney disease
- explain the causes and clinical consequences of selected disorders of water, electrolyte and urate homeostases
- describe how selected disorders of water, electrolyte and urate homeostases are investigated
- discuss the management and treatments of such disorders

8.1 INTRODUCTION

Homeostasis is the maintenance of a stable environment within the body. This stability is necessary for optimum functioning of proteins, particularly enzymes and signalling molecules, cells, tissues, organs and body systems. Many substances have to be maintained at appropriate concentrations, including water, electrolytes (such as Na^+, K^+, Cl^-, Ca^{2+}, Mg^{2+} and P_i) and the acid–base components H^+, CO_2 and HCO_3^- (the last, hydrogen carbonate, is still often referred to as 'bicarbonate'; see *Chapter 9*). In addition to CO_2, other waste products, such as urea and urate, must be kept below toxic levels. Normally, biochemical and physiological mechanisms regulate and control the concentrations of all these components and, in general, homeostatic controls involve negative feedback mechanisms. A receptor detects unacceptable levels of a particular substance under homeostatic control. The receptor then sends a signal to a regulatory centre, which then initiates a response that corrects the imbalance and returns conditions to a physiologically acceptable state (*Figure 8.1*). Once normality returns, the receptor is no longer stimulated and the centre ceases to respond. Disorders of homeostasis can occur, often as a result of failures in the control mechanisms or because of damage to the regulatory centre by external agents.

The kidneys regulate water and electrolyte concentrations, including acid–base balance (*Chapter 9*), and excrete nitrogenous wastes, for example urea and urate. Disorders of kidney functions account for many clinical investigations carried out in hospital laboratories.

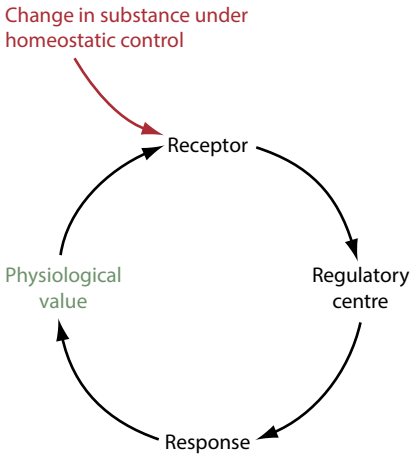


Figure 8.1 A generalised feedback mechanism.

8.2 KIDNEYS

The functions of the kidneys include maintaining the volume of the plasma and its concentrations of electrolytes, such as Na^+ and K^+, and keeping its pH within the normal physiological range. This is largely achieved by varying the amounts of water and salts excreted, the removal of excess H^+ and the regeneration of HCO_3^- (*Chapter 9*). Kidneys also excrete waste products, such as urea, urate and creatinine, and biosynthesise the enzyme renin and the hormones calcitriol (also called 1α,25-dihydroxycholecalciferol [1α,25DHCC]) and erythropoietin (*Figure 8.2*). These control blood pressure, stimulate the production of erythrocytes by the bone marrow (*Chapter 13*) and regulate the

DOI: 10.1201/9781003296836-8

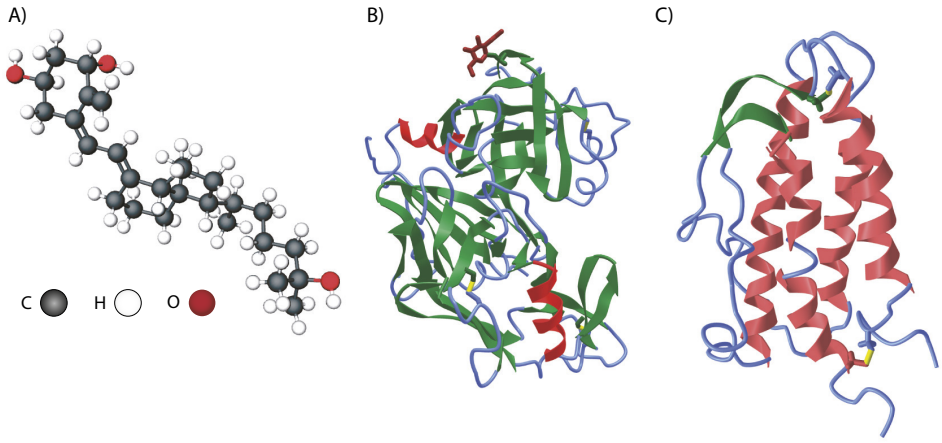


Figure 8.2 (A) Computer drawn model of calcitriol. Molecular models of (B) renin and (C) erythropoietin. PDB files 2V0Z and 1BUY, respectively.

absorption of Ca^{2+} by the gastrointestinal tract (GIT; *Chapter 11*), respectively. Kidneys also biosynthesise prostaglandins and degrade hormones, such as insulin.

Kidney structure

Kidneys are composed of an outer fibrous capsule, a cortex, a middle medulla and an inner pelvis region (*Figure 8.3*).

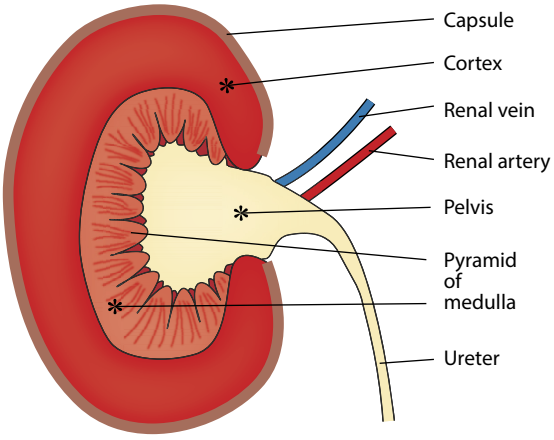


Figure 8.3 Section through a kidney to emphasise its different internal regions. See also *Figure 9.6*.

The tough capsule surrounding each kidney protects against trauma and prevents the entry of bacteria (*Chapter 3*). Kidneys are composed of about a million functional units called **nephrons**, each composed of a tuft of capillaries called a glomerulus and a tubule. Tubules have four different regions, as shown in *Figure 8.4*:

1. a proximal tubule
2. a loop of Henle
3. a distal tubule
4. a collecting duct

The cortex contains the Bowman's capsules and their glomeruli. The medulla is composed of the tubules arranged in pyramids that lead into calyces that excrete urine into the pelvis of the kidney. The pelvis drains into the ureters, which take urine from the kidneys to the bladder. Kidneys have a rich blood supply and normally receive about 25% of the cardiac output. The renal artery supplies blood at a high pressure to afferent arterioles that supply the glomeruli. Glomerular capillaries drain into efferent arterioles that, in turn, divide to form a capillary network covering the nephron (*Figure 8.4*). Blood eventually leaves each kidney in a renal vein.

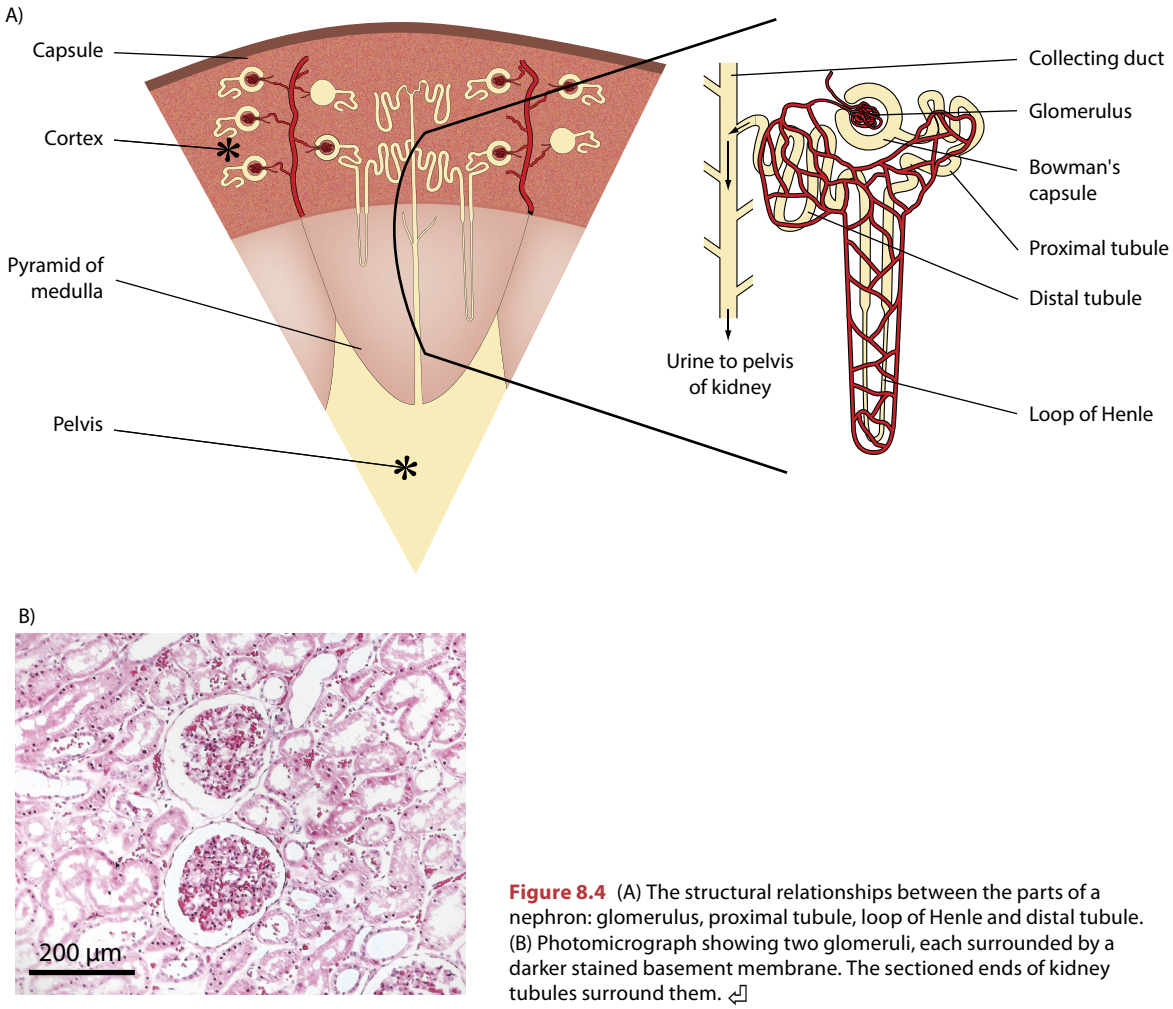


Figure 8.4 (A) The structural relationships between the parts of a nephron: glomerulus, proximal tubule, loop of Henle and distal tubule. (B) Photomicrograph showing two glomeruli, each surrounded by a darker stained basement membrane. The sectioned ends of kidney tubules surround them.

Kidney ultrafiltration and reabsorption

Blood is separated from the lumen of the tubules by three layers: the capillary endothelial cells, a basement membrane and specialised epithelial cells of the Bowman's capsule, called podocytes (*Figure 8.5*). Fenestrae are the openings between the extensions of the podocytes. The basement membrane contains negatively charged glycoproteins that give the basement membrane an overall negative charge. The hydrostatic pressure of blood in the glomerulus is high at 10 kPa because of its direct route from the heart and because the diameters of afferent arterioles supplying the glomeruli are larger than the efferent arterioles collecting blood from them. This forces approximately 20% of the

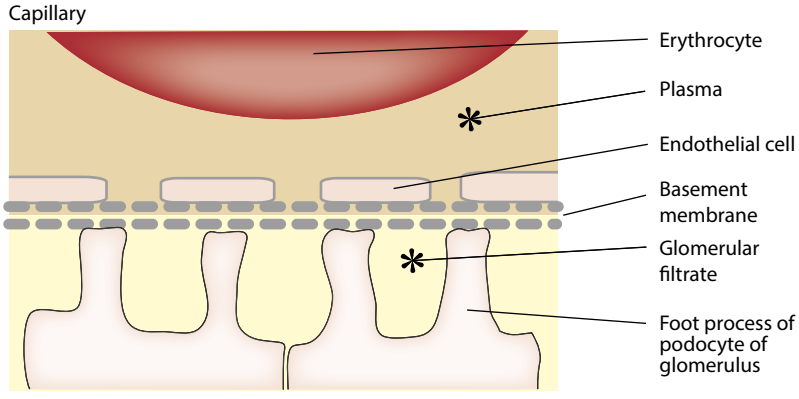


Figure 8.5 Schematic of the filtration unit of the glomerulus.

plasma through the layers into the lumen of the capsule, where it constitutes an ultrafiltrate. The hydrostatic pressure of blood in the glomerulus is opposed by osmotic pressure of 4 kPa generated by its plasma proteins and a back pressure of 2.7 kPa exerted by the filtrate in the Bowman's capsule. Thus, the effective pressure, *Peff*, is:

$$Peff = 10 - (4 + 2.7) = 3.3 \text{ kPa}$$

Water and small molecules are passively filtered into the Bowman's capsule, leaving blood cells, erythrocytes, platelets and plasma proteins in the capillary. Particles with M_r of less than 5000, such as electrolytes, sugars, amino acids, urea and urate, and some small polypeptides and proteins pass freely from the plasma into the lumen of the capsule. Substances with M_r up to 68 000 can penetrate to some extent but larger molecules, such as proteins above 68 000, are excluded because of their size and to some extent their charge, given that most plasma proteins are negatively charged at pH 7.4. Hence, the initial filtrate in the capsule lumen has a composition similar to that of plasma except that it is largely free of protein. Most of the filtrate is reabsorbed as it passes along the nephron. The proximal tubule is responsible for bulk reabsorption of filtrate while the distal tubule is important for fine-tuning its composition depending upon the needs of the body.

Normally, all the glucose, amino acids, K^+ and HCO_3^- and about 75% of the Na^+ are absorbed by energy-dependent mechanisms. The reabsorption of water occurs passively and follows Na^+ reabsorption. Approximately 90% of the filtered Na^+ and 80% of water is reabsorbed in the distal tubule. More Na^+ is reabsorbed in the distal tubule by the cells exchanging it for K^+ and H^+. This exchange is controlled by aldosterone (*Figure 8.6* and *Chapter 7*). Normally, 6–12% of the filtered urate is excreted by the kidneys, with the remainder being reabsorbed in the proximal tubule.

Tubular fluid passes into collecting ducts that extend into the renal medulla and discharge urine into the renal pelvis. About 1–2 L of urine is produced daily, depending on the amount of fluid intake, with larger volumes being produced after an increased intake of water.

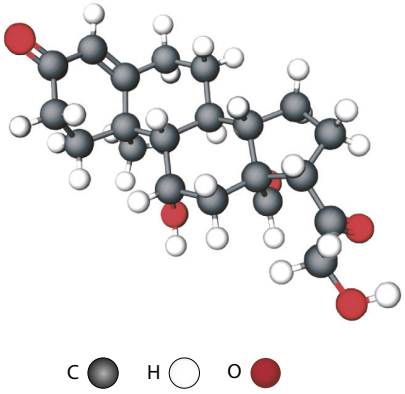


Figure 8.6 Computer drawn model of aldosterone.

Renal function tests

Renal function tests are used to detect the presence of renal diseases and assess their progress. They are, however, of little use in determining the causes of renal disease. The most widely used test is to measure the glomerular filtration rate (GFR), that is, the rate of filtrate formation by the kidneys. The value of the GFR depends on the net pressure across the glomerular membrane, the physical nature of the membrane and its surface area, which reflects the number of functioning glomeruli. All three factors can change because of disease, and this will alter the value of the GFR. In adults, the GFR is about 120 mL/min, although it is related to body size, being higher in men than women. The GFR is also affected by age and declines in the elderly (*Chapter 18*).

Measuring the GFR

The GFR is determined by measuring the concentration of a substance in the urine and plasma that is known to be completely ultrafiltered from the plasma at the glomerulus. This substance must not be reabsorbed or secreted by renal tubules and must remain at a constant concentration in the plasma throughout the period of urine collection. It should also be possible to measure the concentration of this substance in the plasma and urine both conveniently and reliably. Inulin and creatinine have been used to assess the GFR using the equation:

$$\text{GFR} = (U_c \times V) / P_c$$

where U_c is the concentration of substance in urine, P_c is the concentration of substance in plasma and V is the rate of formation of urine in mL/min, meaning that the GFR also has units of mL/min.

Creatinine is derived from creatine phosphate in the muscle, and the amount produced daily is relatively constant. An estimate of creatinine clearance can be made by determining the creatinine concentration in the plasma (see *Figure 1.23*) and the creatinine content in a 24-hour urine collection. Normal creatinine clearance in adults is 115–125 mL/min. However, reliable measurements of creatinine clearance are often difficult because of the need to obtain a complete and accurately timed urine sample.

Measurements of the concentration of serum creatinine may be used to assess renal function, and they are easier to determine than creatinine clearance values. The concentration of creatinine in serum increases with deteriorating renal function but this test lacks sensitivity. For example, the GFR must decrease to less than 50% of the original value before there is a significant increase in serum creatinine. This means that a normal

serum creatinine value does not necessarily exclude the presence of renal disease. An estimated glomerular filtration rate (eGFR) is usually determined from the concentration of creatinine in a serum sample using the formula:

$$eGFR = 175 \times ([\text{serum creatinine}] \times 0.011312)^{-1.154} \times age^{-0.203}$$

where the age of the patient is given in years and the [serum creatinine] in μmol/L. The formula is initially suitable only for White population males and not for females or people of Afro-Caribbean origin, as their renal functions differ. However, it can be applied to females and people of Afro-Caribbean origin by multiplying the final result by 0.742 and 1.21, respectively.

Kidney disease

Kidney disease is the cessation of renal function and it can be acute or chronic. In acute kidney injury (AKI), there is rapid loss of renal function within hours or days, although the condition is potentially reversible and normal renal function can be regained. Clinical deterioration is abrupt, with sudden increases in the concentrations of urea, creatinine and H^+ in serum. Patients with AKI often, but not always, present with oliguria, where there is less than 400 mL of urine passed per day. Indeed, patients are sometimes anuric and fail to pass any urine. Chronic kidney disease (CKD) is the gradual, progressive deterioration of kidney function. As kidney function declines, there is accumulation of waste products that eventually reach toxic levels in the blood and may affect other organs.

Acute kidney injury

Acute kidney injury (previously referred to as acute renal failure) can be categorised as:

- prerenal
- postrenal
- intrinsic

Prerenal AKI can occur because of a decreased plasma volume following blood loss, burns, prolonged diarrhoea or vomiting, decreased cardiac output or occlusion of the renal artery. Whatever the reason, prerenal AKI results in a low GFR and decreased blood flow to kidneys. Aldosterone and antidiuretic hormone (ADH), also called arginine vasopressin (*Figure 8.7*, *Section 8.3* and *Chapter 7*), release is stimulated by the low blood pressure, and the kidneys respond by producing smaller volumes of concentrated urine. Biochemical indicators of prerenal AKI include increased amounts of urea and creatinine in serum due to the low GFR, metabolic acidosis (*Chapter 9*) because of an inability of the kidneys to excrete H^+ and hyperkalaemia (*Section 8.5*) because of the low GFR and acidosis.

Postrenal AKI can be the consequence of a blockage of the urinary tract by, for example, renal calculi, kidney stones or neoplasms. These obstructions increase the hydrostatic pressure opposing glomerular ultrafiltration. If this persists for a sufficiently long time it can cause intrinsic renal damage, which can also result if pre- or postrenal acute injuries are not corrected.

Intrinsic AKI occurs following damage to the kidney itself, which is caused by a variety of conditions. These include nephrotoxins, for example drugs such as aminoglycosides and analgesics, septic shock (*Chapters 2* and *3*), a low cardiac output (*Chapter 14*), burns or crush injuries, and renal diseases such as glomerulonephritis. Glomerulonephritis is inflammation of the renal cortex, which affects the ultrafiltration mechanism of the kidney. It may develop following an infection (*Chapters 2* and *3*).

Three phases occur in AKI:

1. an oliguric phase with a low urine output
2. a diuretic phase where the urine output increases
3. a recovery phase when normal functions return

The oliguric phase is characterised by increased concentrations of K^+, urea, creatinine and H^+ in the serum. If the patient survives the oliguric phase, then urine output increases after a few days when the diuretic phase starts. The GFR increases during the diuretic phase and, as the output of urine increases, the amounts of urea and creatinine in the serum gradually decline. Tubular function may still be abnormal in this phase so that the acidosis may still persist. In the recovery phase, tubular cells regenerate and tubular function is restored to normal. The concentrations of urea and creatinine in serum decrease and K^+ returns to normal concentrations as the GFR improves.

The management of AKI includes correcting prerenal factors if present by, for example, increasing the extracellular fluid volume by administering fluids. Biochemical monitoring of creatinine and K^+ is required and dialysis (see *Figure 12.16*) may be necessary

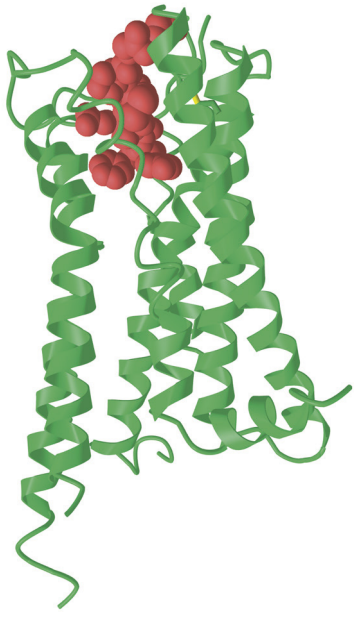


Figure 8.7 Molecular model of antidiuretic hormone (vasopressin), shown in red, bound to a portion of its receptor complex, in green. PDB file 7KH0. ⏎

when K^+ concentrations are high or when severe acidosis is present. The cause of the renal failure should be identified and treated wherever possible.

Chronic kidney disease

Many diseases, such as glomerulonephritis, diabetes mellitus, hypertension and polycystic kidney disease, can lead to irreversible renal damage because they effectively decrease the number of functioning nephrons. Patients may remain asymptomatic until the GFR falls to less than 15 mL/min. Chronic kidney disease (previously called chronic renal failure) progresses to end-stage renal failure (ESRF), where dialysis or a kidney transplant (*Chapter 6*) is necessary for survival. Metabolic features of ESRF include the impairment of urinary concentration and dilution, abnormal electrolyte and H^+ homeostases, the retention of waste products, and decreased biosyntheses of erythropoietin and calcitriol. Chronic kidney disease is accompanied by increases in the concentrations of K^+, P_i, H^+, creatinine and urea, and decreased Ca^{2+} in the plasma.

Patients with CKD present with a number of clinical features including neurological symptoms, such as lethargy, growth failure, myopathy, anorexia, nausea, vomiting, anaemia, hypertension, nocturia and impotence. The causes of many of these are not known but are probably due to the retention of toxins that cannot be excreted. In some cases, it is possible to delay the progression of CKD by treating its cause. A number of measures may be taken to alleviate symptoms before dialysis becomes necessary. These include careful matching of water and Na^+ intakes with their losses. High blood K^+ is controlled with oral ion exchange resins given as their Ca^{2+} or Na^+ salts (*Section 8.4*), whereas a high blood P_i is controlled by oral aluminium or magnesium salts that precipitate ingested phosphate in the GIT. A restriction of dietary protein may reduce the formation of nitrogenous waste.

8.3 DISORDERS OF WATER HOMEOSTASIS

Water is necessary to maintain the volumes of body compartments, to allow the excretion of waste products and as a medium in which biochemical reactions may occur. Water intake is variable and can depend, to some extent, on social habits but is supplied in the diet, from solid foods as well as liquids (*Chapter 10*), and as a product of oxidative metabolism. Its loss is variable, although an almost fixed amount, called the insensible loss, occurs from the GIT, skin and lungs. An average, 70-kg man has 42 L of water distributed between various body compartments (*Figure 8.8*). Water accounts for 60% of body weight in men, but only 55% in women given they have a higher proportion of body fat.

In disease, patients can be **dehydrated**, where water loss caused by vomiting and diarrhoea exceeds gain, or **overhydrated**, with an accumulation of water in body compartments. Clinical features of both are listed in *Table 8.1*. A reduced extracellular fluid (ECF) volume causes a decline in blood circulation with decreased excretion of wastes and reduced oxygen and nutrient supply to the cells. Humans deprived of fluid intake die after a few days because the reduced total body fluid leads to a circulatory collapse.

The kidneys regulate water balance by varying urine output from 0.5 to 15 mL/min to match water intake. When there is an excess of water, the kidneys lose water rapidly, but in times of shortage it is conserved. The total body water is therefore kept constant. Water loss from the kidneys can be regulated by ADH (see *Figure 8.7*). Antidiuretic hormone acts by altering the water permeability of the collecting ducts in the kidneys. Osmoreceptor cells in the hypothalamus detect changes in osmolality (*Box 8.1*) between the intracellular fluid (ICF) and ECF. An increase in the osmolality of the ECF

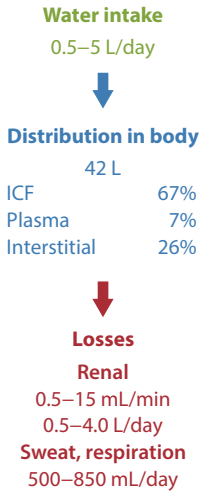


Figure 8.8 The distribution of water in the body. ICF, intracellular fluid.

Clinical feature affected	Dehydration	Overhydration
Pulse	Increased	Normal
Blood pressure	Decreased	Normal/increased
Skin turgor	Decreased	Increased
Eyeballs	Soft/sunken	Normal
Mucous membranes	Dry	Normal
Urine output	Decreased	Normal/decreased
Consciousness	Decreased	Decreased

Table 8.1 Clinical features affected in patients suffering dehydration and overhydration.

BOX 8.1 OSMOLALITY AND THE MOVEMENT OF WATER

Osmolality is the concentration of particles (molecules or ions) dissolved in a fluid. Compartments of cells and, indeed, the body in general are divided by selectively permeable membranes. Most biological membranes are freely permeable to water, which flows by diffusion through transmembrane proteins called aquaporins (*Figure 8.9*). If the osmolality is not the same on both sides of the membrane, then there will be a net movement of water from the side of low osmolality to the one of high, since the osmolality of a body compartment is proportional to the concentration of a compartment expressed as millimoles of solute per kg of water. In humans, the osmolality of serum is approximately 285 mmol/kg. Serum osmolality can be measured directly using an osmometer or, if the concentration of solute is known, it can be calculated in units of mmol/kg using the expression:

$$\text{Osmolality} = 2[Na^+] + 2[K^+] + [\text{glucose}] + [\text{urea}]$$

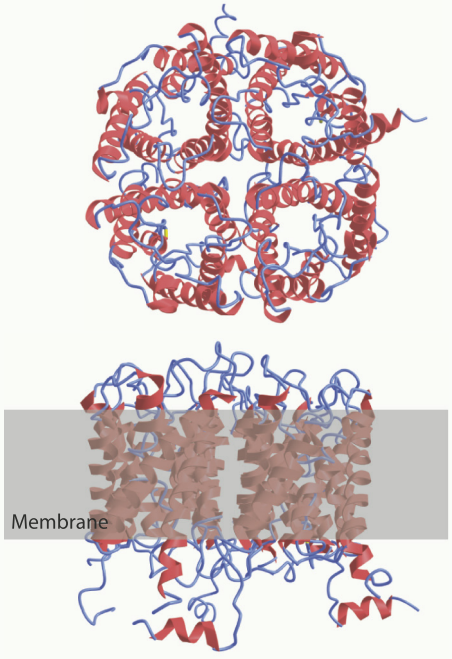


Figure 8.9 Molecular model of aquaporin 2 (AQP2), the major aquaporin of kidney collecting ducts. The grey region shows the approximate position of the membrane. PDB file 4NEF.

stimulates the receptors and these, in turn, stimulate the release of ADH from the posterior pituitary gland (*Chapter 7*). Antidiuretic hormone then stimulates the kidneys to retain water, producing a more concentrated urine. The retention of water helps return the osmolality of the ECF back to normal. If the osmolality of the ECF is low, the osmoreceptors are not stimulated and ADH is not released. This results in water loss from the kidneys in a dilute urine. The loss of water helps to increase the osmolality of the ECF back to normal values. A low blood or ECF volume can be detected by baroreceptors in the aortic arch and carotid sinus (*Chapter 14*). These receptors also stimulate a release of ADH and, indeed, this mechanism can override the release of ADH by osmolality to maintain blood volume and therefore circulation. Interactions between ADH and aldosterone (see *Figures 8.6* and *8.7*) maintain the normal volume and concentration of the ECF. Aldosterone, a steroid hormone, is produced by the adrenal cortex (*Chapter 7*) and released in response to a low ECF volume or blood pressure. It stimulates retention of Na^+ together with water in the kidneys, returning the ECF volume back to normal. There are distinctive signs and symptoms associated with loss of water from body compartments. For example, loss of water from the ICF results in cell dysfunction that presents clinically as confusion, lethargy and coma. Loss of water from the ECF decreases blood pressure, leading to renal shutdown and shock. A reduction in total body water (ICF and ECF) produces a combination of both effects.

The assessment of fluid (and electrolyte) disorders in patients is a significant workload in hospital pathology laboratories. In most cases, clinical tests to determine the concentrations of electrolytes in blood must be interpreted in conjunction with a clinical examination, which involves taking the patient's clinical history, looking for signs and symptoms of hydration or dehydration and assessing kidney function.

8.4 DISORDERS OF ELECTROLYTE HOMEOSTASIS

All body fluids contain electrolytes (*Table 8.2*). The regulation of water content by ADH helps to maintain normal electrolyte concentrations within the body. The concentrations of Na^+ and K^+ in the ICF and ECF are maintained largely by the activity of the plasma membrane Na^+/K^+-ATPase (*Figure 8.10* and *Chapter 11*). This enzyme acts as an energy-dependent pump that expels Na^+ from the cell in exchange for an intake of K^+ to maintain both at physiological concentrations. The concentrations of these ions are maintained within narrow ranges and, since water can flow freely through most

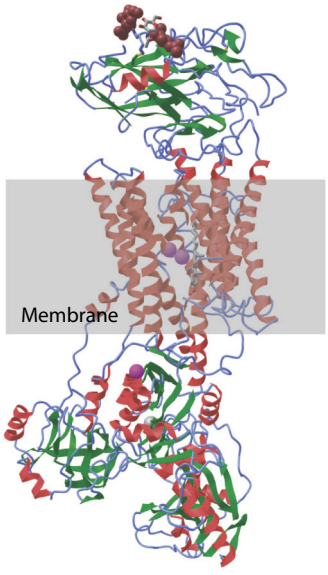


Figure 8.10 Molecular model of a Na^+/K^+-ATPase. The grey block shows the approximate position of the membrane. PDB file 2ZXE.

Electrolyte	[Intracellular fluid] (mmol/L)	[Extracellular fluid] (mmol/L)
K^+	110	4
Na^+	10	135
Cl^-	5	100
HCO_3^-	15	28
PO_4^{2-}	31	1

Table 8.2 Typical compositions of the intracellular fluid (ICF) and extracellular fluid (ECF).

membranes, the concentrations of Na^+ and K^+ are responsible for maintaining the appropriate osmolarities of these compartments. The movement of water from one compartment to another is mainly responsible for determining their volumes. Homeostatic mechanisms exist to minimise changes in body water and electrolyte composition and are particularly important in maintaining the volume of the ECF. Water will remain in the extracellular compartment only if its osmolality is sufficiently high.

Disorders of Na^+ homeostasis

Sodium ions are significant constituents of tissues, including bone; they control the volume of ECF and are required for normal neuromuscular functions. The intake of Na^+ is variable, from less than 100 mmol to more than 300 mmol/day. Losses are also variable, but renal loss is normally matched to intake. Small amounts of Na^+ are lost via skin and in faeces and, under some circumstances, the GIT can be a major route of Na^+ loss, as in diarrhoea (*Chapter 11*).

The average, 70-kg man contains 3700 mmol of Na^+ (*Figure 8.11*), of which 75% is found in the ECF. *Hyponatraemia* and *hypernatraemia* refer to serum concentrations of Na^+ below and above the reference range of 135–145 mmol/L, respectively. Hyponatraemia is caused by an excessive retention of water or the loss of Na^+, with these two conditions resulting in different clinical features. The retention of water produces behavioural disturbances, headaches, confusion, convulsions and eventually coma. The symptoms associated with excessive loss of Na^+ are weakness, apathy, dizziness, weight loss and hypotension. Hyponatraemia due to water retention is the more common.

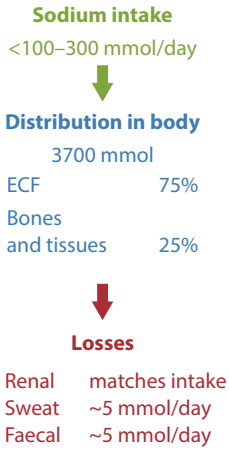


Figure 8.11 The distribution of sodium in the body. ECF, extracellular fluid.

Oedema

Oedema is the excessive accumulation of fluid in interstitial compartments. It may occur with or without an increase in total body Na^+. The former produces an oedematous hyponatraemia, whereas the latter results in nonoedematous hyponatraemia. The excessive accumulation of water is a result of increases in the concentration of Na^+ in the ECF, which gives rise to swollen tissues. These may be localised in, for example, legs and ankles (*Figure 8.12*) but can be more generalised, occurring in the chest cavity, abdomen and lungs. The major causes of oedematous hyponatraemia are heart failure (*Chapter 14*), liver disease (*Chapter 11*) and nephrotic syndrome (*Margin Note 8.1*). All three reduce blood volume and stimulate aldosterone secretion, which, in turn, results in the retention of Na^+. The reduced blood volume also promotes the release of ADH from the posterior pituitary. Both result in more water than Na^+ being retained, giving rise to hyponatraemia. Treatments for oedematous hyponatraemia are aimed at its underlying

Margin Note 8.1
Nephrotic syndrome

Nephrotic syndrome leads to a loss of blood proteins to the urine, with the reduced concentrations of plasma albumin leading to oedema. Its most common cause is renal damage by diseases such as glomerulonephritis.

Figure 8.12 A massive oedema of the lower limb. Courtesy of Charlie Goldberg, MD, medicine.ucsd.edu/clinicalmed.

cause, for example heart failure or kidney or liver diseases, and at removing the excess water and Na^+ using diuretics, and restricting water intake.

Nonoedematous hyponatraemia, the result of water overload without an increase in total body Na^+, is due to a decreased excretion of water from the syndrome of inappropriate secretion of ADH (SIADH), a severe kidney injury or an increased intake by compulsive drinking or excessive parenteral fluid. The condition SIADH is a common finding in clinical practice. Patients present with reduced plasma osmolality, normal kidney function but a low output of urine. The syndrome is associated with many conditions, including malignancies, for example carcinoma of the lungs or bowel (*Chapter 17*); infections, such as pneumonia and tuberculosis (*Chapter 3*); and trauma following, for example, abdominal surgery; or it may be induced with drugs, such as chlorpropamide (*Figure 8.13*). All these conditions result in SIADH with water retention and a low urinary output. The excessive water is distributed between ICF and ECF and so the clinical signs of water overload, oedema, may be mild or absent. Treatment of SIADH is to reduce water intake to less than 750 mL/day and to correct its underlying cause.

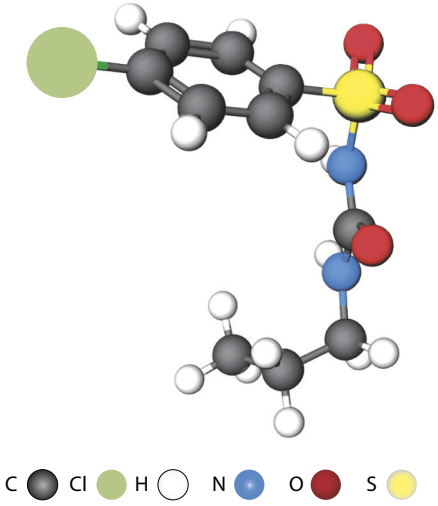


Figure 8.13 Computer drawn model of chlorpropamide.

Hyponatraemia

Hyponatraemia from loss of Na^+ decreases the total body Na^+ of patients. The losses may occur from vomiting, diarrhoea, kidney dysfunction (aldosterone deficiency), the effects of drugs such as spironolactone (*Figure 8.14*), or a decreased dietary intake of Na^+, although this is very rare. The loss of Na^+ is always accompanied by water loss; as the volume of the ECF decreases, the release of ADH is stimulated and the increased reabsorption of water produces hyponatraemia. The decreased volume of ECF means that the patient presents with the clinical symptoms of dehydration. Treatment is aimed at correcting the Na^+ losses with intravenous infusions of 0.9% NaCl and treating the underlying cause, for example steroid therapy for aldosterone deficiency as in Addison disease (*Chapter 7*).

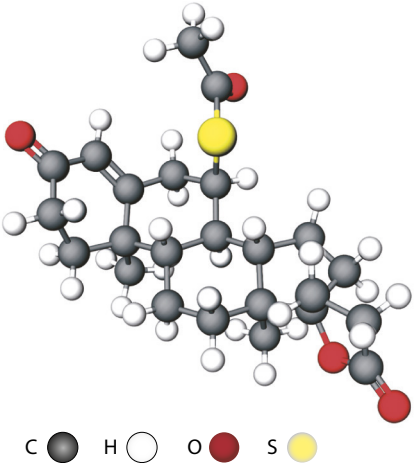
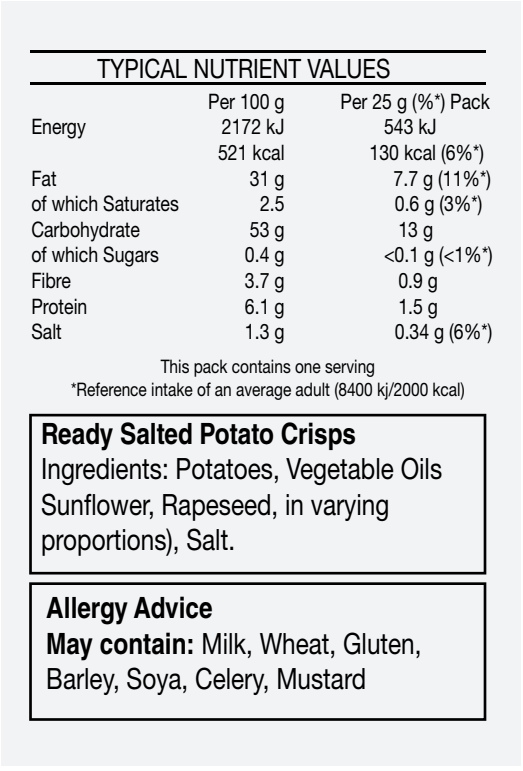


Figure 8.14 Computer drawn model of spironolactone.

Hypernatraemia

Hypernatraemia is caused by water depletion, water and Na^+ depletion with the loss of water predominating, or an excess of Na^+ (*Figure 8.15*). Its clinical features are variable but, in general, patients present with muscular weakness, hypertension, intense thirst (**polydipsia**) and polyuria. If fluid loss occurs, the features associated with dehydration may be present. However, if Na^+ is in excess, raised blood pressure or oedema may be seen. Water depletion results from a decreased intake, such as in comatose patients, infants or the elderly. The body conserves water by producing a low volume of concentrated urine. Increased water losses can also occur in diseases such as diabetes insipidus that result in large quantities of dilute urine (*Chapter 7*). Hypernatraemia with water and Na^+ depletion occurs only if relatively more water than Na^+ is lost. It is commonly caused in children by excessive sweating or diarrhoea. Patients respond by producing low volumes of concentrated urine. The condition may also occur during osmotic diuresis in patients with diabetes mellitus (*Chapter 7*) where both water and Na^+ are lost together with other electrolytes in large volumes of dilute urine, producing hypernatraemia and a decreased ECF. An excess of Na^+ in the ECF is caused by an increased intake or decreased excretion of Na^+. The intake may be oral, for example salt tablets or seawater, or parenteral as in the treatment of Conn and Cushing syndromes (*Chapter 7*). Both disorders produce dilute urine due to retention of Na^+ by the kidneys.



TYPICAL NUTRIENT VALUES

	Per 100 g	Per 25 g (%*) Pack
Energy	2172 kJ 521 kcal	543 kJ 130 kcal (6%*)
Fat	31 g	7.7 g (11%*)
of which Saturates	2.5	0.6 g (3%*)
Carbohydrate	53 g	13 g
of which Sugars	0.4 g	<0.1 g (<1%*)
Fibre	3.7 g	0.9 g
Protein	6.1 g	1.5 g
Salt	1.3 g	0.34 g (6%*)

This pack contains one serving
*Reference intake of an average adult (8400 kj/2000 kcal)

Ready Salted Potato Crisps
Ingredients: Potatoes, Vegetable Oils Sunflower, Rapeseed, in varying proportions), Salt.

Allergy Advice
May contain: Milk, Wheat, Gluten, Barley, Soya, Celery, Mustard

Figure 8.15 Many highly processed foods and snacks, such as potato crisps, contain high levels of Na^+ as its chloride salt (NaCl). ⏎

Hypernatraemia is treated by oral administration of water. If this is not possible, then 5% dextrose is administered parenterally. If hypernatraemia is due to an excessive Na^+ intake, measures to remove it must be considered.

Disorders of K^+ homeostasis

Potassium ions are necessary to maintain cell volume, for the optimal activities of a number of enzymes, and to maintain the resting potential of cell membranes and therefore neuromuscular functions, especially in the heart (*Chapter 14*). The intake of K^+ varies between 30 and 100 mmol/day and losses are equally variable. The kidneys excrete most ingested K^+, with a smaller amount being eliminated by the GIT. A high concentration of plasma K^+ stimulates the release of aldosterone (*Chapter 7*) that, in turn, increases the renal excretion of K^+. Gastrointestinal losses can be significant during vomiting and diarrhoea. Only very small amounts of K^+ are lost in sweat. The average 70-kg human contains about 3600 mmol of K^+ (*Figure 8.16*), almost all being found in the ICF. Values

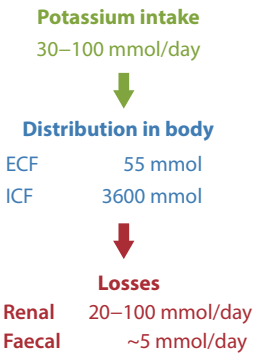


Figure 8.16 The distribution of potassium in the body. ECF, extracellular fluid; ICF, intracellular fluid. ⏎

for the concentration of K^+ in the serum below and above the reference range of 3.4–4.9 mmol/L are called *hypokalaemia* and *hyperkalaemia*, respectively. Hyperkalaemia is the more common clinical condition.

Hypokalaemia

Hypokalaemia is clinically significant, giving rise to muscular weakness and cardiac arrhythmias (*Chapter 14*); hence, patients often present with breathlessness and chest pain. The causes of hypokalaemia include increased K^+ losses from the GIT or kidneys, alkaloses, certain clinical disorders, some drugs or a decreased K^+ intake. Excessive losses from the GIT can occur during vomiting and diarrhoea. Hypokalaemia occurs in alkalosis because the pH of the ECF is high and H^+ moves from the ICF to the ECF as part of the buffering process, while K^+ moves in the opposite direction causing an hypokalaemia. A number of disorders, for example Cushing and Conn syndromes (*Chapter 7*), are associated with increased cortisol and aldosterone production, respectively. Both hormones have mineralocorticoid activity and stimulate the renal retention of Na^+ in exchange for K^+ leading to hypokalaemia. Drugs, such as carbenoxolone (see *Figure 9.11*) used to treat gastric ulcers (*Chapter 11*), can cause hypokalaemia because of their mineralocorticoid activity. Decreases in oral or parenteral intakes of K^+ are rare but can produce an hypokalaemia. Patients with hypokalaemia are treated with oral K^+ salts. Severe hypokalaemia may require intravenous infusions of K^+.

Hyperkalaemia

Hyperkalaemia must be identified because concentrations of serum K^+ above 7 mmol/L can result in cardiac arrest and death, even though, unlike Na^+, the plasma K^+ concentration does not vary significantly with water loss or overload. Renal failure, acidoses, aldosterone deficiency, damage to cells and an excessive intake of K^+ can all cause hyperkalaemia. In renal failure, the kidneys are unable to excrete K^+ because of the low GFR. Further, acidosis, a common feature of kidney disease, leads to hyperkalaemia because the low pH of the ECF means that K^+ moves out of cells in exchange for H^+, to return the pH to reference values. A deficiency of aldosterone, such as in Addison disease where the kidneys lose their ability to excrete K^+, can result in hyperkalaemia. The destruction of cells during trauma can release large amounts of K^+ causing hyperkalaemia. Lastly, an excessive oral or parenteral intake of K^+ is a rare cause of hyperkalaemia. The treatment of hyperkalaemia includes infusion of insulin and glucose to promote the entry of K^+ into cells. Severe hyperkalaemia may require dialysis.

Disorders of Ca^{2+} homeostasis

Calcium is required for bone and teeth structure (see *Boxes 10.4* and *11.2*), the release of neurotransmitters and initiation of muscle contraction, as a cofactor for coagulation factors (*Chapter 13*) and some enzyme activities, and it also acts as an intracellular second messenger for a number of hormones (*Chapter 7*).

The normal dietary intake of Ca^{2+} of about 25 mmol/day is supplemented by the reabsorption of Ca^{2+} from GIT secretions. Approximately 19 mmol of Ca^{2+} is lost in the faeces daily. The kidneys normally filter about 240 mmol of Ca^{2+} daily but, as most of this is reabsorbed by the tubules, normal renal loss of Ca^{2+} is only about 6 mmol/day. Calcium is the most abundant mineral in the body and the average adult contains approximately 1 kg or 25 000 mmol of Ca^{2+} (*Figure 8.17*). Approximately 99% of Ca^{2+} is present in the bone, of which about 500 mmol is exchanged daily with the ECF. The ECF contains about 22.5 mmol of Ca^{2+}, of which 9.0 mmol is present in the plasma. Approximately 47% of Ca^{2+} in plasma occurs as free, ionised Ca^{2+}, 46% is protein bound and 7% is complexed with citrate or phosphate. Only free Ca^{2+} is physiologically active and its plasma concentration is controlled by homeostatic mechanisms involving the hormones parathyroid hormone (PTH), calcitriol and calcitonin (*Figure 8.18*). Parathyroid hormone is secreted by the parathyroid glands in response to a fall in the concentration of plasma ionised Ca^{2+} and vice versa. It stimulates **bone resorption**, that is, the release of Ca^{2+} from bone, and a decreased reabsorption of HCO_3^- by the kidneys that produces an acidosis, which helps to increase plasma ionised Ca^{2+} and stimulates the synthesis of calcitriol from cholecalciferol in the liver. This hormone is also formed in the skin by the action of ultraviolet light on 7-dehydrocholesterol (see *Figure 10.20*). Calcitriol increases Ca^{2+} and P_i absorption from the GIT and increases bone resorption. The physiological function of calcitonin remains unclear, but it is known to reduce the concentration of Ca^{2+} in plasma by inhibiting both bone resorption and reabsorption by the kidneys. The serum reference range for total Ca^{2+} is 2.20–2.60 mmol/L and for free is 1.20–1.37 mmol/L. Values below and above these are called *hypocalcaemia* and *hypercalcaemia*, respectively.

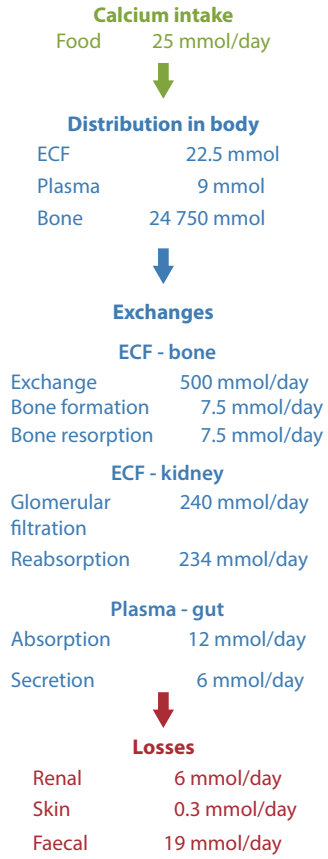


Figure 8.17 The distribution of calcium in the body. ECF, extracellular fluid.

Hypocalcaemia

The clinical effects of hypocalcaemia include behavioural disturbances, paraesthesiae, tetany, convulsions and cataracts. Its major causes are kidney disease, Mg^{2+} and vitamin D deficiencies, hypoparathyroidism and pseudohypoparathyroidism. Chronic kidney disease may decrease the reabsorption of Ca^{2+} by decreasing the biosynthesis of calcitriol leading to hypocalcaemia. This may lead to bone disease because the increased output of PTH arising from the hypocalcaemia can increase osteoclast activity. Magnesium ions are required for PTH secretion and its action and a deficiency produces hypocalcaemia. A deficiency in vitamin D may arise from a poor diet, malabsorption (*Chapters 10* and *11*) or inadequate exposure to sunlight with an inadequate absorption of Ca^{2+} from food. Hypoparathyroidism, or a reduced activity of the parathyroid glands with decreased production of PTH, results in hypocalcaemia. The condition can be congenital, where there is an absence of parathyroid glands, or acquired hypoparathyroidism that may be **idiopathic**, or caused by autoimmune conditions or surgery, for example thyroidectomy. In pseudohypoparathyroidism, there is excessive PTH secretion because target tissues fail to respond to the hormone, producing a persistent hypocalcaemia. This condition is more common in males than females and patients present with skeletal abnormalities including short stature, intellectual disabilities, cataracts and testicular atrophy.

An algorithm for investigating hypocalcaemia is outlined in *Figure 8.19*. The underlying cause of hypocalcaemia should be treated wherever possible. Oral Ca^{2+} supplements

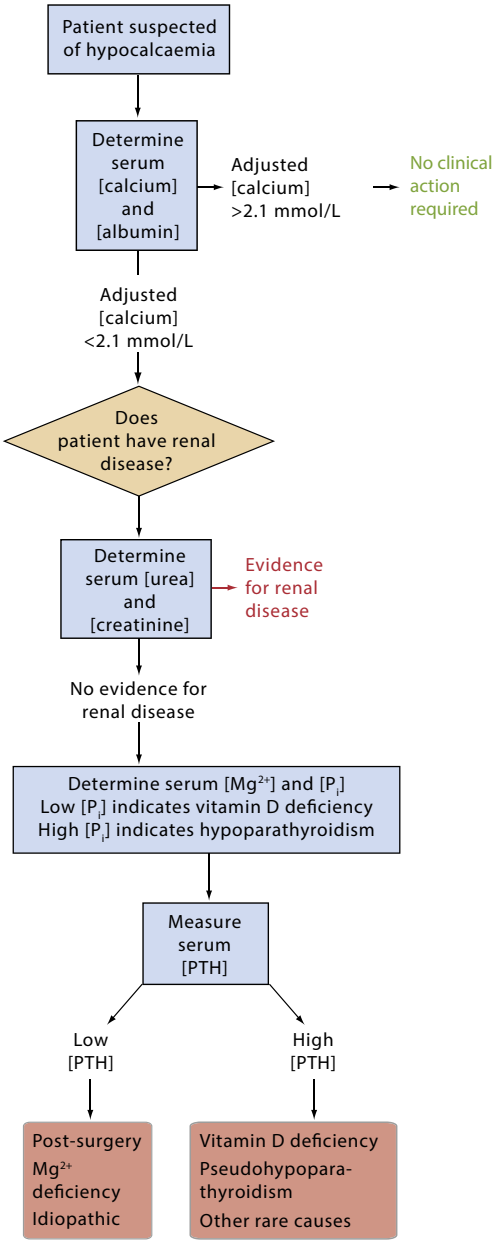


Figure 8.19 Overview of the clinical investigation of hypocalcaemia. PTH, parathyroid hormone.

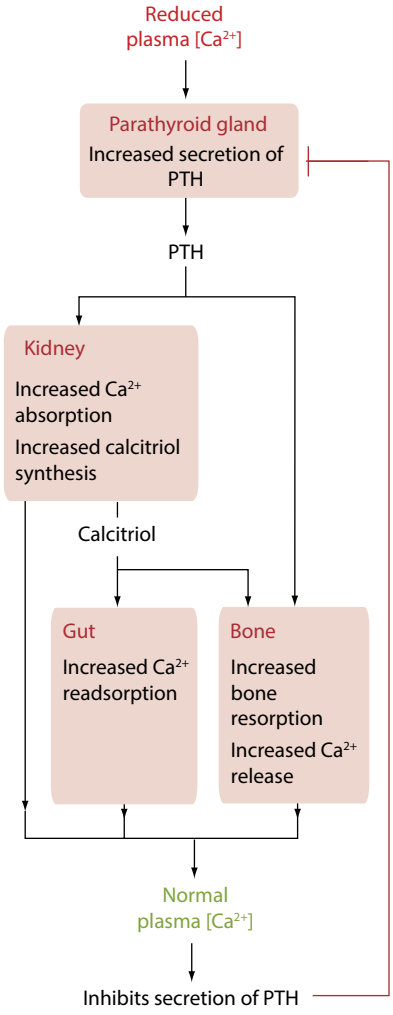


Figure 8.18 The hormonal regulation of the concentration of plasma Ca^{2+}. PTH, parathyroid hormone.

are prescribed in mild cases of hypocalcaemia. Magnesium supplements may be prescribed in hypocalcaemia due to Mg^{2+} deficiency, whereas calcitriol and its precursors may be prescribed in vitamin D deficiency.

Hypercalcaemia

Renal damage is the most serious consequence of hypercalcaemia. It may also suppress neuromuscular excitability causing constipation and abdominal pain and affect the central nervous system (CNS), resulting in depression, nausea and anorexia. The nausea may cause vomiting and therefore dehydration. Calcium can stimulate gastrin and therefore gastric acid secretion and so hypercalcaemia may be associated with peptic ulcers (*Chapter 11*). Hypercalcaemia may cause arrhythmias and in severe cases may result in cardiac arrest (*Chapter 14*).

The most common causes of hypercalcaemia are malignant disease or primary hyperparathyroidism. Cancerous tumours of the lungs stimulate an increase in plasma Ca^{2+} by producing a PTH-related protein (PTHrp) that resembles it in structure (*Figure 8.20*). Cytokines and prostaglandins released by tumours that have metastasised to the bones may lead to increased resorption of Ca^{2}. Primary hyperparathyroidism occurs most commonly due to a parathyroid adenoma, which is a benign tumour, and only rarely due to a parathyroid carcinoma. It affects both men and women at any age but is most common in postmenopausal women. In primary hyperparathyroidism, there is excessive PTH secretion that causes hypercalcaemia and sometimes hypophosphataemia (see below), which increases bone turnover, particularly of the metaphyses (*Figure 8.21*).

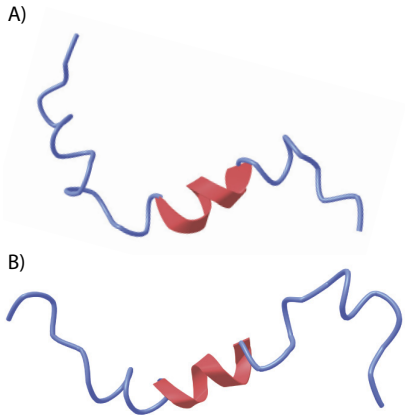


Figure 8.20 Molecular models of (A) parathyroid hormone (PTH) and (B) parathyroid hormone-related protein (PTHrp). PDB files 1HTH and 1BZG, respectively. Note the similarity in both conformations.

Figure 8.21 An isotope scan of hand and finger bones showing sites in the fingers with a decreased opacity where greater bone turnover has occurred because of increased Ca^{2+} resorption. Courtesy of Dr I. Maddison, London South Bank University, UK.

Margin Note 8.2
Familial hypocalciuric hypercalcaemia

Familial hypocalciuric hypercalcaemia is an autosomal dominant (*Chapter 15*) condition that develops from childhood. It is characterised by chronic hypercalcaemia but is usually asymptomatic, with normal concentrations of PTH and no parathyroid adenoma. Familial hypocalciuric hypercalcaemia is caused by mutations in the human calcium sensing receptor, encoded by *CASR* sited at 3q13.33-q21.1.

Thyroid hormones have no direct effect on Ca^{2+} homeostasis but can cause increased bone **iatrogenic** turnover by increasing osteoclast activity and giving rise to mild hypercalcaemia during thyrotoxicosis. An excessive iatrogenic or accidental ingestion of vitamin D or, much less commonly, thiazide diuretics (see *Figure 9.11*) that interferes with renal Ca^{2+} loss can also lead to hypercalcaemia. Less common causes of hypercalcaemia include thyrotoxicosis, vitamin D intoxication (*Chapter 10*) and familial hypocalciuric hypercalcaemia. Rare causes are tuberculosis, sarcoidosis, acromegaly, milk-alkali syndrome and idiopathic hypercalcaemia of infancy.

Both sarcoidosis and tuberculosis are granulomatous diseases. In these conditions, hypercalcaemia occurs, as there is increased production of calcitriol by macrophages in the granulomas. Hypercalcaemia is occasionally seen in acromegaly, probably due to stimulation of calcitriol production by excess growth hormone. Hypercalcaemia may occur in people who ingest large amounts of milk together with alkali antacids, such as HCO_3^-, to relieve symptoms of peptic ulceration. An alkalosis occurs that is believed to reduce renal Ca^{2+} excretion although the precise mechanism is still unclear. Milk-alkali syndrome is rare, as antacid treatment of peptic ulcers has been replaced by treatments that eliminate *Helicobacter pylori* and which also inhibit gastric acid secretion (*Chapters 3* and *11*).

The condition idiopathic hypercalcaemia of infancy is associated with hypercalcaemia because of an increased sensitivity to vitamin D in bone and the GIT. It occurs as two types, that is, infantile hypercalcaemia 1 and infantile hypercalcaemia 2, which are distinguishable by mutations in different genes: type 1 in *HCINF1* and type 2 in *SLC34A1* and *CYP24A1*.

Patients who present with hypercalcaemia are investigated for malignancy or primary hyperparathyroidism, as these account for up to 90% of clinical cases. If both malignancy and primary hyperparathyroidism are excluded, other causes must be considered and investigated (*Figure 8.22* and *Box 8.2*). A number of approaches are taken to managing hypercalcaemia.

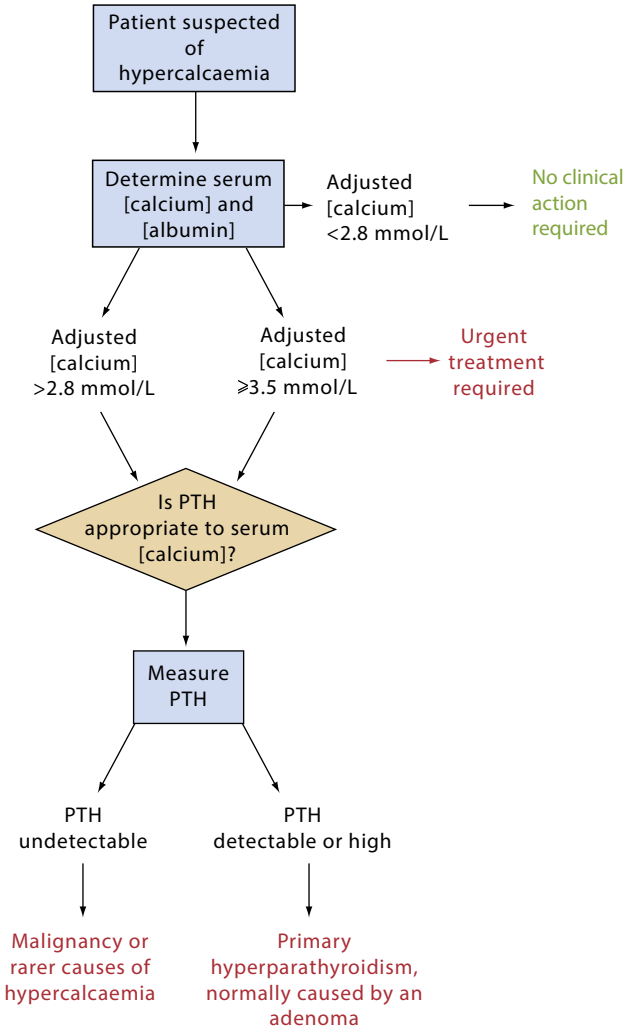


Figure 8.22 Overview of the clinical investigation of hypercalcaemia. See also *Box 8.2*. PTH, parathyroid hormone.

The underlying cause should be treated wherever possible. Intravenous saline may be administered to dehydrated patients to restore the GFR, enhance Ca^{2+} loss and rehydrate the patient. Drugs, such as frusemide, inhibit renal reabsorption of Ca^{2+} and promote its excretion, while bisphosphonates (*Figure 8.25*) lower Ca^{2+} concentrations by inhibiting bone resorption. In extreme cases, dialysis or emergency parathyroidectomy may be necessary.

In some cases, an artefactual hypocalcaemia may be reported when blood samples are erroneously collected into tubes containing ethylenediaminetetraacetic acid (EDTA). This anticoagulant is a chelator of Ca^{2+} (and other metal cations [see *Figure 12.28*]) and its use will lead to low reported values for Ca^{2+} concentrations.

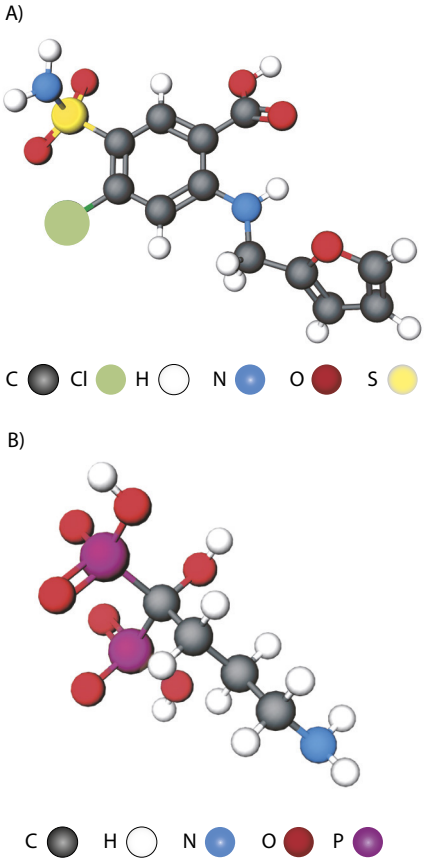


Figure 8.25 Computer drawn models of (A) frusemide and (B) the bisphosphonate pamidronate.

BOX 8.2 ALBUMIN AND Ca^{2+}

Changes in the concentration of albumin in plasma (*Figure 8.23*) affect the total Ca^{2+} concentration but not the free Ca^{2+} content. A high plasma albumin concentration gives rise to a high total plasma Ca^{2+} concentration and, conversely, a low plasma albumin concentration reduces the total plasma Ca^{2+} concentration (*Figure 8.23*). The determination of total plasma or serum Ca^{2+} concentration can be misleading, since it is affected by conditions that change the concentration of albumin. This drawback is overcome by measuring the concentration of free Ca^{2+}, which is now the standard practice. The adjusted value for total Ca^{2+} can be calculated using the following formulae if the [albumin] is given in g/L and $[Ca^{2+}]$ in mmol/L.

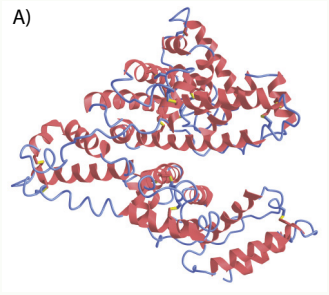
When [albumin] is <40 g/L then:

adjusted $[Ca^{2+}]$ = measured concentration of total Ca^{2+} + 0.02(40 – [albumin])

When [albumin] is >45 g/L then:

adjusted $[Ca^{2+}]$ = measured concentration of total Ca^{2+} – 0.02([albumin] – 45)

Since the free Ca^{2+} concentration is unaffected by variations in the concentration of albumin, clinical symptoms are not manifested. However, free Ca^{2+} competes with H^+ for negatively charged binding sites on albumin, and changes in the concentration of free Ca^{2+} may occur in acute acid–base disorders, with clinical consequences (*Chapter 9* and *Figure 8.24*). These do not, however, affect the total plasma Ca^{2+} concentration. In alkaloses, more H^+ dissociates from albumin, allowing increased amounts of free Ca^{2+} to bind to it. This has the effect of increasing the protein bound Ca^{2+} fraction but at the expense of decreasing the free plasma Ca^{2+} leading to hypocalcaemia. In acidoses, increasing amounts of H^+ bind to albumin as Ca^{2+} dissociates from it, decreasing the protein bound Ca^{2+} and increasing free plasma Ca^{2+} producing hypercalcaemia.



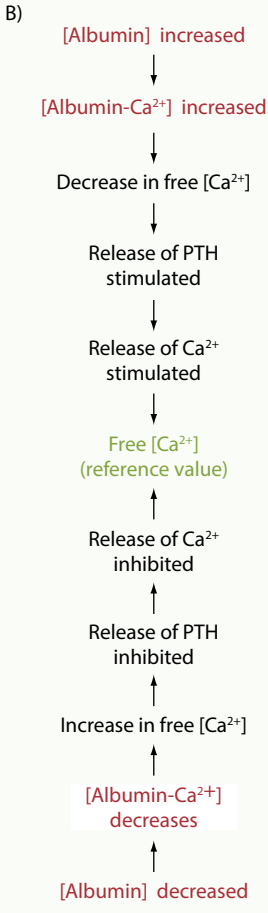


Figure 8.23 (A) Molecular model of human serum albumin. PDB file 1E78. (B) The effect of changes in the concentration of albumin on the concentration of total plasma Ca^{2+}. PTH, parathyroid hormone.

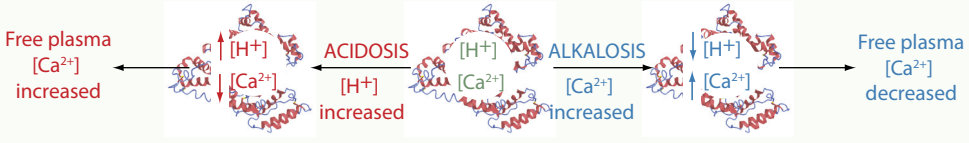


Figure 8.24 Effect of acidosis and alkalosis on the concentration of free Ca^{2+} in the plasma.

Disorders of P_i homeostasis

Phosphate (P_i) combines with Ca^{2+} to form hydroxyapatite, the mineral component of bone and teeth (see *Boxes 10.4* and *11.2*). It is also required for some enzymic activities, oxidative phosphorylation, the biosynthesis of 2,3-bisphosphoglycerate that regulates the dissociation of oxyhaemoglobin (*Chapter 13*), the excretion of H^+ (*Chapter 9*) and cell membrane integrity. The daily intake of P_i is about 40 mmol. The kidneys lose approximately 26 mmol daily and 14 mmol are lost in faeces. The total body content of P_i in the average male is over 20 000 mmol (*Figure 8.26*) with 17 000 occurring in bone and 3000 in soft tissues, largely attached to lipids and proteins. Thus, about 85% occurs in bone while the ICF and the ECF contain 15% and 0.1%, respectively. The plasma concentration is about 1 mmol/L. Approximately 80% of the plasma content occurs as free P_i, 15% is protein bound and about 5% is complexed with Ca^{2+} and Mg^{2+}. Parathyroid hormone and calcitriol (see *Figures 8.20* and *8.2*) control the homeostasis of P_i, the former decreasing the reabsorption by the kidneys and reducing its plasma concentration and the latter stimulating P_i absorption in the GIT and increasing the concentration.

The reference range for total serum P_i is 0.8–1.4 mmol/L but a higher reference range applies in infancy and childhood. *Hypophosphataemia* and *hyperphosphataemia* are used to describe concentrations below and above the reference range, respectively. Hypophosphataemia causes more damage than hyperphosphataemia but, fortunately, is less common.

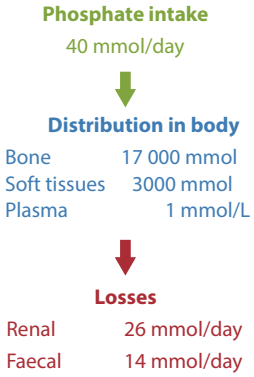


Figure 8.26 The distribution of phosphate in the body.

Hypophosphataemia

Clinical features of hypophosphataemia include paraesthesiae, ataxia, coma, osteomalacia and muscle weakness. There may be increased susceptibility to infection possibly due to defective phagocytosis (*Chapter 4*). The causes of hypophosphataemia are varied. Vitamin D deficiency results in a decreased biosynthesis of calcitriol and therefore decreased P_i absorption by the GIT. Increased renal loss of P_i may occur in primary hyperparathyroidism where increased secretion of PTH causes excessive renal loss of P_i. Certain diuretics that increase renal loss of P_i can cause hypophosphataemia. It may also occur during the recovery phase of diabetic ketoacidosis when patients are administered insulin, which promotes cellular uptake of P_i. Total body P_i may be depleted as a consequence of osmotic diuresis. There are a number of rare causes of hypophosphataemia, which include an inadequate dietary intake usually associated with parenteral nutrition; when agents, such as aluminium hydroxide, are used as antacids and prevent its absorption in the GIT; and in chronic alcoholics who have a complex and multifactorial condition with poor diet and reduced GIT absorption (*Chapter 11*).

Determination of the serum concentrations of P_i and Ca^{2+} and the urinary concentration of P_i are useful in investigating hypophosphataemia. The following strategy may be used when its cause is not obvious. First, exclude causes such as alkalosis and chronic alcoholism. Second, a reduced urinary P_i suggests decreased dietary or parenteral intakes or increased cellular uptake, for example in insulin therapy. Third, if the urinary concentration of P_i is above its reference range then excessive renal losses are occurring and the concentration of Ca^{2+} in the plasma or serum should be determined. If this is increased, then primary hyperparathyroidism or malignancy may be present. If, however, the concentration is normal or low, renal defects or inappropriate diuretic therapy are considerations. Hypophosphataemia should be managed by treating the underlying cause wherever possible. In some situations it may be necessary to administer oral or parenteral P_i.

Hyperphosphataemia

Hyperphosphataemia may cause metastatic calcification, for example the deposition of calcium phosphate in soft tissues, as the excess P_i precipitates with Ca^{2+} and causes hypocalcaemia and tetany in affected patients. The most common cause of hyperphosphataemia is renal failure where the GFR and P_i excretion decline. Hypoparathyroidism reduces renal excretion of P_i giving rise to hyperphosphataemia. In diabetic ketoacidosis (*Chapter 7*), a deficiency of insulin prevents the uptake of P_i by cells leading to hyperphosphataemia. Other causes are an increased intake of P_i or its release from damaged cells in intravascular haemolysis. Indeed, any condition where there is increased turnover of cells, for example following treatment of malignant disease with chemotherapy, results in release of P_i during cell destruction. Excessive intake, either oral or intravenous, is a rare cause and is more likely when there is also renal failure as in pseudohypoparathyroidism where there is resistance by the kidneys to PTH that decreases their excretion of P_i.

A number of biochemical tests are useful when investigating hyperphosphataemia. These include determining the concentrations of P_i, Ca^{2+}, urea and creatinine in serum and the concentration of P_i in urine. The following strategy has proved useful in investigating obscure causes of hyperphosphataemia. First, it is necessary to exclude

artefactual causes. For example, a delay in the separation of plasma or serum from blood before analysis for P_i or haemolysis of a blood sample prior to its analysis can produce an artefactual hyperphosphataemia that does not reflect the true clinical situation. Second, serum concentrations of creatinine and urea should be determined to exclude renal failure. If the serum concentration of Ca^{2+} is normal or above reference values, vitamin D intoxication or untreated diabetes mellitus should be considered. Third, if the plasma or serum concentration of Ca^{2+} is low, then hypoparathyroidism should be investigated. Finally, if the urinary concentration of P_i is low, then hypoparathyroidism is, again, a consideration, whereas a high urinary concentration indicates increased intake, malignancy or intravascular haemolysis.

Patients with hyperphosphataemia are managed by treating the underlying cause wherever possible. The oral intake of aluminium, Ca^{2+} and Mg^{2+} salts may be used, as these can bind P_i in the GIT reducing its absorption.

Disorders of Mg^{2+} homeostasis

Magnesium ions are required to maintain the structures of ribosomes, nucleic acids and numerous proteins and act as a cofactor and for a variety of enzymes, including those involved in energy metabolism and protein synthesis, often as the salt of nucleotides in these cases (*Figure 8.27*). It is also required for normal cell permeability and neuromuscular functions.

The usual dietary intake of Mg^{2+} is about 15 mmol/day and approximately 30% of this is absorbed in the GIT; the rest is lost in the faeces. The adult human body contains over 1200 mmol of Mg^{2+} (*Figure 8.28*). Approximately 750 mmol are found in bone and about 450 mmol in muscle and soft tissues. The ECF contains only 15 mmol. Approximately 55% of plasma Mg^{2+} occurs as free ionised Mg^{2+}, 32% is protein bound and 13% is complexed with P_i or citrate.

The kidneys lose 5–10 mmol of Mg^{2+} daily but losses are adjusted to control Mg^{2+} homeostasis. An increased dietary intake of Mg^{2+} results in increased renal loss and vice versa. This is achieved principally by adjusting the reabsorption of Mg^{2+} by cells of the proximal tubules and loop of Henle. A number of factors influence the rate of excretion of Mg^{2+}, including hypercalcaemia and hypophosphataemia (see above) that decrease renal reabsorption and PTH, which stimulates renal retention.

The reference range for serum Mg^{2+} is 0.8–1.2 mmol/L. *Hypo-* and *hypermagnesaemia* refer to concentrations below and above the reference range, respectively. Note that measurements of the concentration of Mg^{2+} in plasma or serum are unreliable indicators of its body status, since only 1% of body Mg^{2+} occurs in the ECF.

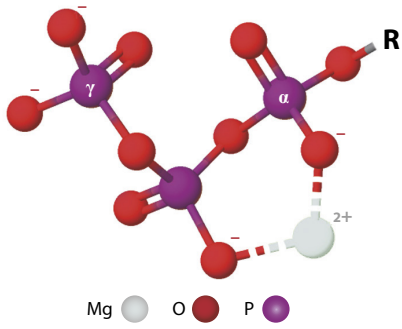


Figure 8.27 The β,γ complex form of MgATP, which is the favoured structure in near neutral solutions. R, adenosine. The complex is a cofactor for many enzymes.

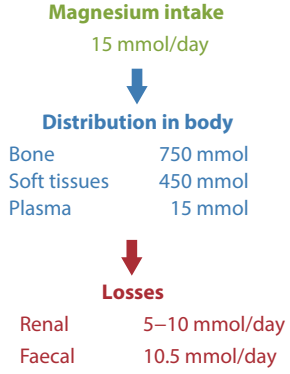


Figure 8.28 The distribution of magnesium in the body.

Hypomagnesaemia

Hypomagnesaemia is believed to occur due to increased renal excretion, inadequate dietary intake, vomiting and diarrhoea. Its clinical effects are similar to those seen in hypocalcaemia (see above) and include tetany, muscle weakness, convulsions and cardiac arrhythmias. These effects are related to the role of Mg^{2+} in neuromuscular function. Its causes include decreased intake as in starvation (*Chapter 10*), poorly managed parenteral nutrition or malabsorption. Increased losses of Mg^{2+} as in osmotic diuresis in diabetics (*Chapter 7*), diuretic therapy, hyperaldosteronism and excessive losses from the GIT in prolonged diarrhoea, GIT fistula and laxative abuse can also result in hypomagnesaemia. The use of anticancer drugs (*Chapter 17*), such as cisplatinum, can damage the kidneys and prevent the renal reabsorption of Mg^{2+}. In alcoholism (*Chapter 12*), hypomagnesaemia is believed to occur due to increased renal excretion, inadequate dietary intake, vomiting and diarrhoea.

In many cases, the cause of hypomagnesaemia is determined by clinical examination. However, measuring the urinary Mg^{2+} may be useful, as the amount of Mg^{2+} excreted per day decreases with decreased intake. If hypomagnesaemia occurs with increased renal excretion then losses are likely to be due to renal damage. Hypercalcaemia may increase renal Mg^{2+} excretion causing hypomagnesaemia, but hypocalcaemia may occur in hypomagnesaemia due to hypoparathyroidism.

The underlying cause of hypomagnesaemia should be treated wherever possible. Oral Mg^{2+} supplements may be adequate for mild cases but severe Mg^{2+} deficiency together with malabsorption may require intravenous infusions of Mg^{2+}.

Hypermagnesaemia

The most common causes of hypermagnesaemia are AKI or CKD; less common are its release from damaged cells from, for example, crush injuries. Mild hypermagnesaemia may occur in mineralocorticoid deficiency, as in Addison disease. In rare cases, hypermagnesaemia may occur from an increased oral or parenteral intake of Mg^{2+} or from the use

of Mg^{2+} containing antacids or laxatives. When this does occur, it is usually combined with renal failure. The clinical effects of hypermagnesaemia are also largely related to the role of Mg^{2+} in neuromuscular activities and include muscular weakness, respiratory paralysis and, in extreme cases, cardiac arrest. Its management involves treating the underlying cause wherever possible. Hypermagnesaemia due to renal failure may require dialysis.

8.5 DISORDERS OF URATE METABOLISM

The end product of the catabolism of purines, adenine and guanine is urate (*Figure 8.29*). There are three sources of purines, namely:

1. diet
2. the catabolism of endogenous nucleotides and nucleic acids
3. *de novo* biosynthesis

Most dietary nucleic acids are ingested in the form of nucleoproteins from which urate is produced by the GIT (*Chapter 11*). The catabolism and *de novo* synthesis of purines are linked (*Figure 8.30*). The body's urate pool, and therefore plasma concentration, depends upon the relative rates of urate formation and excretion. Both the kidneys and the GIT excrete urate, with renal excretion accounting for approximately 66% of the total. Almost all the urate is filtered at the glomerulus, but most is reabsorbed by the proximal tubule. However, both reabsorption and secretion occur in the distal tubule, so that the net effect is to excrete about 10% of the urate. Urate secreted into the GIT is degraded to CO_2 and NH_3 by bacterial action or **uricolysis**.

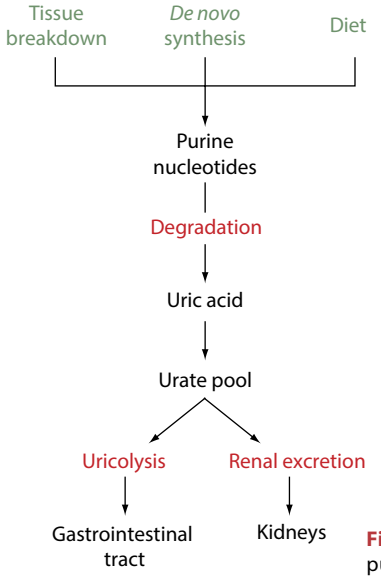


Figure 8.30 The sources and catabolism of purine bases. See also *Figure 6.31*.

The reference range for serum urate is 0.1–0.4 mmol/L. Values below this are referred to as *hypouricaemia*. Even in health there is a wide variation in the concentration of urate in plasma. Plasma urate concentration tends to be higher in males than females and is highest in obese individuals, those from affluent social classes and those with high protein and alcohol intakes. Hence, *hyperuricaemia* is defined as a concentration greater than 0.42 mmol/L in men and more than 0.36 mmol/L in women.

Hypouricaemia

Hypouricaemia is uncommon and generally not of clinical significance. Its occurrence is due to a decreased urate synthesis, as in congenital xanthine oxidase deficiency (*Figure 8.31*) and severe liver disease, or to increased excretion of urate as seen in renal tubular disorders, such as Fanconi syndrome. Hypouricaemia may also result from excessive use of drugs such as allopurinol (see *Figure 8.33*).

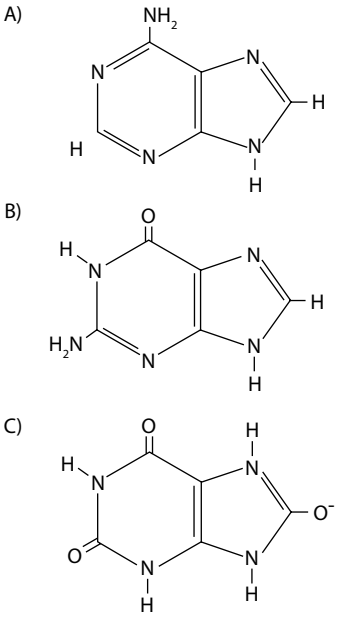


Figure 8.29 (A) Adenine, (B) guanine and (C) urate.

Margin Note 8.3
Uric or urate?

Uric acid in its native state is a weak acid, that is, it is an *-ic*. However, in the buffered, near-neutral pH of most biological fluids, it is ionised and occurs as its salt, that is, it is an *-ate*.

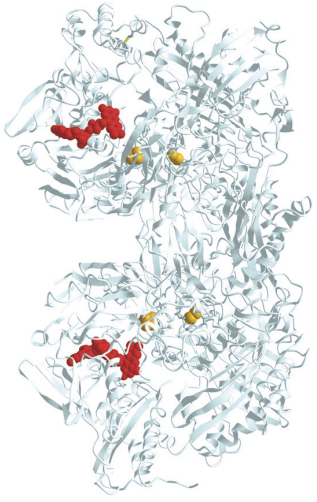


Figure 8.31 Molecular model of xanthine oxidase, with its bound flavin adenine dinucleotide (FAD) and FeS centres shown in red and orange, respectively. PDB file 2CKJ.

Hyperuricaemia

Hyperuricaemia may arise as a result of increased production of uric acid or decreased excretion or, indeed, both. Excessive synthesis may occur because of a defective biosynthetic metabolic pathway, stimulation of *de novo* purine production by alcohol, increased nucleic acid turnover, as in malignant disease, or the use of cytotoxic drugs. An excessive dietary intake of purines will also produce hyperuricaemia. A decreased urate excretion may be due to a reduced GFR giving rise to hyperuricaemia. Increased proximal tubular reabsorption and decreased distal tubular secretion of urate have similar effects. Lactate and β hydroxybutyrate compete with urate for excretion by the distal tubule. Therefore, lactic acidosis or ketosis (*Chapter 7*) are often associated with hyperuricaemia. Some drugs, for example low doses of aspirin, can inhibit the distal tubular secretion of urate causing hyperuricaemia.

Urate has low solubility and the ECF easily becomes saturated at concentrations just above the upper limit of the reference range. There is a tendency for crystalline monosodium urate to form in people with hyperuricaemia. Crystals of monosodium urate, which are needle shaped and 2–10 μm long, tend to form in cartilage and synovial fluid of joints, particularly those of the big toe, causing gout (*Figure 8.32*). The crystals are phagocytosed by neutrophil leukocytes and may cause damage to lysosomal membranes within these cells (*Chapter 16*). As a consequence, lysosomal contents are released, causing damage to both leukocytes and surrounding tissues with an associated inflammatory response (*Chapter 4*).

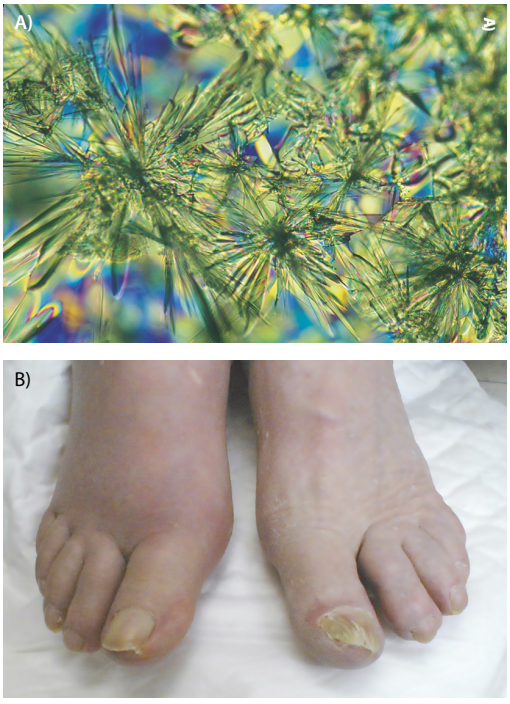


Figure 8.32 (A) Crystals of monosodium urate viewed using polarised light. (B) Gout of the right big toe showing diffuse swelling and inflammation centred where the toes join the foot but also extending over much of the foot. Courtesy of Charlie Goldberg MD, medicine.ucsd.edu/clinicalmed. ⏎

Gout may be primary, with no known cause, or secondary, as a consequence of another disorder. Primary gout is characterised by recurrent attacks of arthritis. It is more common in men than women. The metabolic defect in patients is unknown but a number of abnormalities may be responsible for the overproduction of urate and therefore increased urinary urate output. In many patients, there is a combined defect of urate overproduction together with its impaired renal excretion. Patients with primary gout often have deposits of urate in their soft tissues, and some can develop renal stones composed of urate salts. The risk that a normal person will develop gout varies with their urate concentration. The annual incidence of gout in men is low, at about 0.1% when the urate concentration is less than 0.42 mmol/L. This increases to 0.6% when the concentration is 0.42–0.54 mmol/L and 5% when urate concentrations are greater than 0.54 mmol/L. The reason for the onset of acute attacks in gout is unclear, since a sharp rise in the concentration of urate is not usually demonstrable.

Secondary gout is rare but can arise from a number of other disorders including myeloproliferative disorders (*Chapter 17*) such as polycythaemia vera, where the hyperuricaemia is due to an increased cell turnover, the use of cytotoxic drug therapy that

increases cell destruction and the breakdown of nucleic acids, and psoriasis with its increased turnover of skin cells.

A diagnosis of gout is made on clinical grounds, a demonstration of hyperuricaemia and a satisfactory response to uricosuric drugs. A high plasma urate concentration does not always mean that the patient has gout, but a high concentration of urate in the plasma makes the condition more likely, whereas a consistently low plasma urate concentration excludes the diagnosis. To confirm a preliminary diagnosis, it is necessary to aspirate the joint fluid during an acute attack. The finding of urate crystals in neutrophils will confirm the diagnosis.

Nonsteroidal anti-inflammatory drugs, such as colchicine, ibuprofen and indomethacin, are used to treat acute attacks of gout but have no effect on the hyperuricaemia, which is treated with a diet low in protein and alcohol. Urate-lowering drugs, for example allopurinol (*Figure 8.33*), which prevent the formation of urate and decrease *de novo* synthesis of purines, are used in long-term treatment or when plasma urate levels are persistently higher than 0.6 mmol/L.

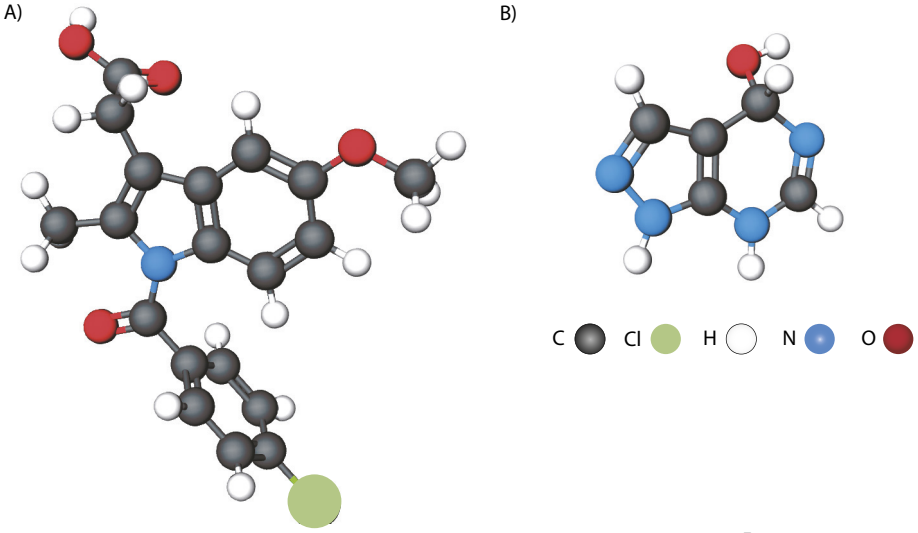


Figure 8.33 Computer drawn structures of (A) indomethacin and (B) allopurinol.

8.6 CASE STUDIES

CASE STUDY 8.1

Ted was admitted to hospital following a car accident. The following results were obtained on a serum specimen three days later (reference ranges are given in parentheses):

Urea	45	(2.4–6.5 mmol/L)
Na^+	132	(133–145 mmol/L)
K^+	6.9	(3.4–4.8 mmol/L)
HCO_3^-	13	(21–28 mmol/L)
Osmolality	332	(280–290 mmol/kg)

Questions

(a) What may be the primary cause of these results?

(b) Suggest reasons why Ted shows hyperkalaemia and low serum HCO_3^-.

CASE STUDY 8.2

Arnie, a 25-year-old man, presented with a history of severe diarrhoea, abdominal pain, weight loss, cramp in the arms and legs, and tetany. He had suffered several previous episodes of diarrhoea and abdominal pain. His serum was investigated and yielded the following results (reference ranges are given in parentheses):

Na^+	140	(133–145 mmol/L)
K^+	3.3	(3.4–4.8 mmol/L)
Urea	5.8	(2.4–6.5 mmol/L)
Total Ca^{2+}	2.40	(2.15–2.46 mmol/L)
Phosphate	1.0	(0.8–1.44 mmol/L)
Albumin	42	(38–48 g/L)
Mg^{2+}	0.39	(0.8–1.2 mmol/L)

Questions

(a) Explain the significance of these results.
(b) Suggest ways in which Arnie should be treated.

CASE STUDY 8.3

Zeshan, a 62-year-old accountant who has had bowel cancer, was admitted to hospital when feeling unwell. A routine blood sample gave the following test results (reference ranges are given in parentheses):

Na^+	120	(133–145 mmol/L)
K^+	3.8	(3.4–4.8 mmol/L)
Urea	2.3	(2.4–6.5 mmol/L)
Creatinine	53	(70–150 µmol/L)
Glucose	4.0	(2.8–6.0 mmol/L)
Osmolality	236	(282–296 mmol/kg)

Questions

(a) What is the most plausible diagnosis for Zeshan?
(b) Explain how the clinical data support your diagnosis.

CASE STUDY 8.4

John, a 58-year-old obese lecturer, woke in the middle of the night with severe pain in his large toe, which was hot, swollen and red. The pain was so intense he could not place his foot on the floor. John had been to a dinner party the night before. In the morning, he visited the local hospital where a blood sample was taken and analysed for serum urate, giving the following result (reference range in parentheses):

Serum urate	0.81	(0.1–0.4 mmol/L)

Questions

(a) What is the most likely diagnosis?
(b) What further investigations should be performed?
(c) How should John be treated?

8.7 SUMMARY

It is essential to maintain appropriate levels of water and electrolytes in the body so that metabolic reactions can function effectively. The stable environment within the body is maintained by homeostatic mechanisms, which return concentrations to normal following shifts in equilibria. In addition, waste products such as urate need to be removed to prevent toxicity. The kidneys help to maintain the balance of water, electrolytes and waste products, and a number of renal function tests are available to assess their function in cases of suspected renal failure. Dehydration, possibly as a result of gastrointestinal disease causing diarrhoea and vomiting, is a severe, life-threatening condition. Disorders of electrolyte balance can involve a lack or excess of the electrolyte in question. Thus, distinct syndromes can occur with disorders affecting the concentrations of Na^+, K^+, Cl^-, Ca^{2+}, P_i and Mg^{2+}. Disorders of urate metabolism may result in high levels of urate in the blood, leading to gout and renal stones, whereas low levels of it in the blood are rare.

QUESTIONS

1. Robin was rescued from a raft at sea. He had been without food or water for several days.
   (a) What will have happened to Robin's body compartments?
   (b) Should he have drunk seawater to survive?

2. Jane, an 85-year-old woman who lives alone, was suffering from a urinary tract infection and had little food or water for several days. She was found in a drowsy confused state by her neighbours and taken to hospital. An analysis of her serum gave Na^+ and urea concentrations of 160 and 20 mmol/L, respectively. Glucose and K^+ concentrations were within their references ranges. Account for Jane's symptoms and test results.

3. Hyperkalaemia may be caused by which ONE of the following?
   (a) Haemolysis.
   (b) Delayed separation of plasma.
   (c) Increased intake of K^+ supplements.
   (d) Kidney failure.
   (e) All of the above.

4. What is the adjusted serum Ca^{2+} concentration for a patient with a serum albumin concentration of 29 g/L and a total serum Ca^{2+} of 1.78 mmol/L?

5. Which ONE of the following does NOT cause hyperuricaemia?
   (a) Excessive dietary purine intake.
   (b) Malignant disease.
   (c) CKD.
   (d) Hypomagnesaemia.
   (e) Psoriasis.

6. Alice, a 48-year-old woman, was treated by parathyroidectomy for her hyperparathyroidism. Unfortunately, she developed hypocalcaemia and was placed on vitamin D therapy, although the hypocalcaemia proved difficult to control. Eventually, Alice was seen as an outpatient complaining of feeling unwell and vomiting. Her blood results showed her to be hypercalcaemic but other values were within their reference ranges. Suggest the most plausible reason for Alice's hypercalcaemia.

7. Explain why hypercalcaemia is often associated with malignancy.

8. Indicate which of the following statements are TRUE and which are FALSE.
   (a) The kidneys regulate water and electrolyte concentrations but not acid–base balance, which is purely a function of the lungs.
   (b) The enzymes renin, calcitriol and erythropoietin help control blood pressure, stimulate the production of erythrocytes by the bone marrow and regulate the absorption of Ca^{2+} by the GIT.
   (c) The cortex of the kidney contains the Bowman's capsules and their glomeruli, and the medulla is composed of kidney tubules arranged in pyramids.
   (d) In the kidney, a basement membrane supports the capillary endothelial cells and specialised cells called podocytes.
   (e) Prerenal kidney disease can occur if blood loss leads to a decreased plasma volume.

(f) Postrenal kidney disease by obstructions of the urinary tract can reduce the rate of glomerular ultrafiltration.
(g) A number of diseases and clinical conditions can irreversibly damage the kidneys and effectively decrease the number of functioning nephrons.
(h) Dehydration is associated with an increased pulse but overhydration has the opposite effect of reducing the pulse.
(i) Aquaporin 4 (AQP4) is the major aquaporin found in kidney collecting ducts.
(j) Nephrotic syndrome leads to a loss of blood proteins, including albumin, to the urine.

9. Arrange the following two columns into their most appropriate pairings:

Water	Daily intake of approximately 40 mmol
Dehydration	Hypokalaemia
Na^+	Syndrome of inappropriate secretion of ADH
Na^+/K^+-ATPase	Pulse increased
Nonoedematous hyponatraemia	Idiopathic hypercalcaemia of infancy
Carbenoxolone	Appropriate intra- and extracellular osmolarities
Ca^{2+}	Most abundant of body minerals
Hypercalcaemia	60% of body weight in men
P_i	Mild hypermagnesaemia
Addison disease	Intra- and extracellular concentrations of 10 and 135 mmol/L, respectively

10. The pK_a of the ionisable H^+ of uric acid is 5.8. Use the Henderson–Hasselbalch equation (see *Box 9.1*) to calculate the proportion of acid to its salt (urate) at the physiological pH 7.4.

FURTHER READING

Bernal A *et al*. (2023) Sodium homeostasis, a balance necessary for life. *Nutrients* **15**:395. doi: 10.3390/nu15020395.

Charles C and Ferris AH (2020) Chronic kidney disease. *Prim Care* **47**:585–595.

Clebak KT *et al*. (2020) Gout: rapid evidence review. *Am Fam Physician* **102**:533–538.

Fiorentini D *et al*. (2021) Magnesium: biochemistry, nutrition, detection, and social impact of diseases linked to its deficiency. *Nutrients* **13**:1136. doi: 10.3390/nu13041136.

Hunter RW and Bailey MA (2019) Hyperkalemia: pathophysiology, risk factors and consequences. *Nephrol Dial Transplant* **34**(suppl. 3):iii2–iii11.

Lindner G *et al*. (2022) Hyponatremia in the emergency department. *Am J Emerg Med* **60**:1–8.

Palmer BF and Clegg DJ (2016) Physiology and pathophysiology of potassium homeostasis. *Adv Physiol Educ* **40**:480–490.

Pepe J *et al*. (2020) Diagnosis and management of hypocalcemia. *Endocrine* **69**:485–495.

Pickkers P *et al*. (2021) Acute kidney injury in the critically ill: an updated review on pathophysiology, and management. *Intensive Care Med* **47**:835–850.

Tonon CR *et al*. (2022) A review of current clinical concepts in the pathophysiology, etiology, diagnosis, and management of hypercalcemia. *Med Sci Monit* **28**:e935821.

9

Disorders of acid–base balance

OBJECTIVES

After studying this chapter you should be able to:

- describe how hydrogen ions (H^+) are transported by the blood
- outline the role and operations of blood buffering systems
- calculate the concentrations of weak acids and their salts necessary to prepare buffered solutions of defined pH and concentration
- explain the role of kidney tubule cells in excreting H^+ and regenerating plasma hydrogen carbonate (HCO_3^-, 'bicarbonate')
- explain the causes of disorders of acid–base balance
- describe the strategies used to investigate disorders of acid–base balance
- outline how disorders of acid–base balance are clinically managed

9.1 INTRODUCTION

The concentration of hydrogen ions ($[H^+]$) in the blood is kept within a narrow reference range to give the blood a pH of approximately 7.4. The body possesses biochemical and physiological mechanisms that maintain this pH by removing excess H^+ and CO_2 produced during metabolism (*Figure 9.1*). These activities are vital for normal bodily functions and are performed by the renal and respiratory systems, respectively. Failures to maintain an appropriate acid–base balance give rise to an **acidosis**, with a blood pH below the reference range, or an **alkalosis**, with a pH above it. Different types of acidoses and alkaloses produce specific characteristic clinical features. Once a specific acid–base disorder has been identified, a clinical strategy must then be adopted to manage the symptoms and to treat the underlying cause(s).

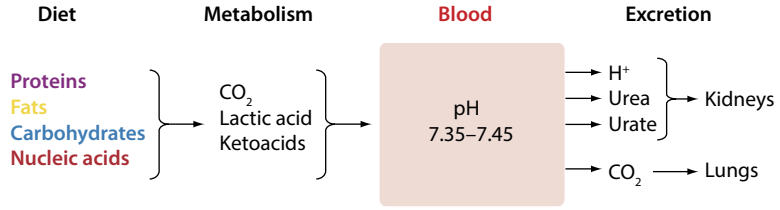


Figure 9.1 Overview of the production, transport and excretion of CO_2 and H^+.

9.2 THE PRODUCTION AND TRANSPORT OF CO_2

Body tissues produce approximately 20 moles of CO_2 per day during oxidative metabolism. The CO_2 diffuses from the cells into the extracellular fluid (ECF), that is, tissue fluid, and then blood, reaching quantities with the potential to form enough carbonic acid (H_2CO_3) to disturb its pH. However, in normal circumstances, this does not occur because the CO_2 is transported to the lungs where it is excreted. During transport, a substantial proportion of the CO_2 enters the erythrocytes by diffusion. Within the erythrocytes, a small proportion of the CO_2 remains dissolved or combines with proteins, mainly haemoglobin, to form carbamino compounds:

DOI: 10.1201/9781003296836-9

$$\text{Protein-NH}_2 + \text{CO}_2 \rightleftharpoons \text{Protein-NH-COO}^- + \text{H}^+$$

The major portion, however, combines with water to produce H_2CO_3 in a reaction catalysed by carbonic anhydrase (*Figure 9.2*). Carbonic acid dissociates to H^+ and HCO_3^-

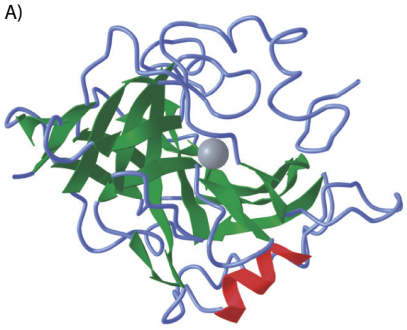


Figure 9.2 (A) Molecular model of carbonic anhydrase. The grey sphere represents a Zn^{2+} in its active site. PDB file 2CBD. (B) The reaction catalysed by carbonic anhydrase.

(hydrogen carbonate, often referred to by the outmoded term 'bicarbonate'):

$$H_2CO_3 \rightleftharpoons H^+ + HCO_3^-$$

Figure 9.3 shows how H^+ are removed from solution when they react with oxyhaemoglobin (HbO_8), which promotes the release of its O_2 to the tissues and forms protonated haemoglobin ('H^+Hb'). The HCO_3^- formed diffuses down its electrochemical gradient out of the erythrocytes to the plasma in exchange for Cl^-, thus maintaining the electrochemical equilibrium within the erythrocyte. The exchange of HCO_3^- for Cl^- is normally

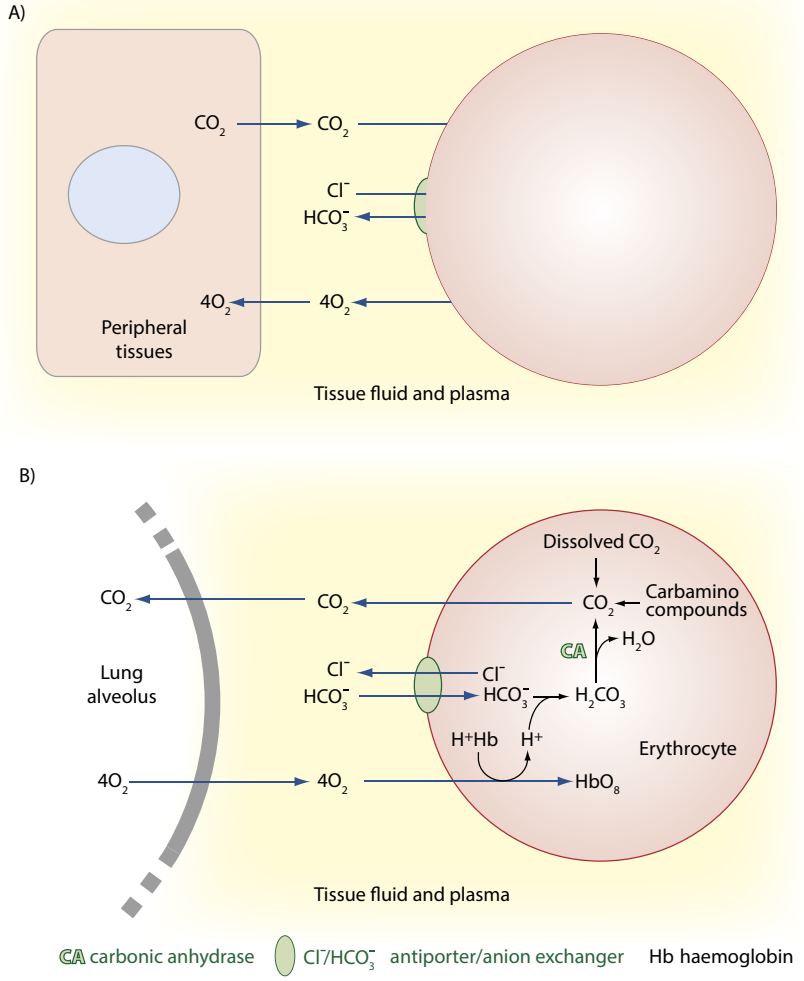


Figure 9.3 (A) The uptake of CO_2 by an erythrocyte with the release of O_2. (B) The excretion of CO_2 and uptake of O_2 at the lung epithelium. Note the Cl^- shift in both cases.

called the **chloride shift**. Since both ions are charged, neither would pass freely across biological membranes. However, an anion exchanger membrane protein facilitates their cotransport through the membrane in opposite directions: hence, the exchanger is referred to as an antiporter. The concentration of HCO_3^- ($[HCO_3^-]$) in the plasma is normally kept between 21 and 28 mmol/L.

In the lungs, the partial pressure of O_2 is high while that of CO_2 is low. Thus, O_2 enters the erythrocytes forming oxyhaemoglobin, releasing the bound H^+ and promoting the reverse of the events that occur in other body tissues (*Figure 9.3*). Thus, H^+ associates with HCO_3^- to produce H_2CO_3, which is broken down to CO_2 and water in a reaction catalysed by carbonic anhydrase. The water enters the large body pool while the CO_2 leaves the erythrocytes and is excreted on exhalation.

Margin Note 9.1
Which direction?

These activities of carbonic anhydrase are an interesting confirmation that enzymes catalyse reactions in either direction depending upon the position of the reaction relative to that of its equilibrium. Thus, carbonic anhydrase promotes the formation of H_2CO_3 in most body tissues where the concentration of CO_2 ($[CO_2]$) is relatively high. However, in the lungs, where the $[CO_2]$ is reduced, the enzyme catalyses the formation of CO_2 and H_2O from H_2CO_3.

9.3 BUFFERING AND THE EXCRETION OF H^+

About 60 mmol of H^+ is produced each day from the oxidation of sulphur-containing amino acid residues or from incomplete metabolic activities, such as anaerobic glucose metabolism or ketone body formation (*Figure 9.4* and *Chapter 7*). If all the H^+ were released into the approximately 14 L of ECF, the $[H^+]$ would be 4 mmol/L or about 100 000 times more acidic than normal. In reality, the [H] is kept within the narrow limits of 40±5 nmol/L by temporary buffering systems that maintain the physiological pH of 7.4±0.05 necessary for normal metabolic functions. These systems include the capture of H^+ by haemoglobin in erythrocytes, as described in the previous section, and the $HCO_3^- - H_2CO_3$ buffer system (*Box 9.1*), which resist changes to the pH of the plasma until the excessive H^+ is excreted by the kidneys (*Chapter 8*). If the concentrations of H^+ and HCO_3^- reach equilibrium, however, buffering would become ineffective. In the case of the $HCO_3^- - H_2CO_3$ system, this is usually prevented by the breakdown of H_2CO_3 to CO_2 and water. The formation of H_2CO_3 from H^+ and HCO_3^- is a rapid reaction. Its potentially slow breakdown to CO_2 and H_2O is accelerated by carbonic anhydrase activity in the erythrocytes and kidneys, and the removal of CO_2 at the lungs prevents the system from reaching equilibrium.

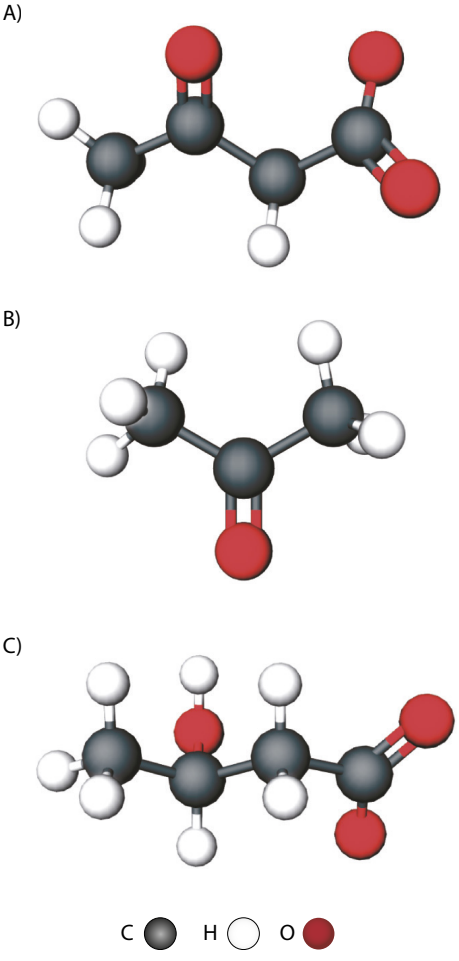


Figure 9.4 Computer drawn models of the ketone bodies (A) acetoacetate, (B) acetone and (C) β hydroxybutyrate.

BOX 9.1 THE RELATIONSHIP BETWEEN H^+, PCO_2 AND HCO_3^-

The pH of a solution is defined as:

$$pH = -\log[H^+]$$

The pH scale (*Figure 9.5*) ranges from 0 to 14 and describes $[H^+]$ of 10^0 (or 1) to 10^{-14} mol/L. A pH of 7 is neutral; values below this are increasingly acidic and those above it increase in alkalinity.

A buffered solution is one that resists changes to its pH when relatively small amounts of acid or alkali are added to it. In organisms, the most significant buffers are protein molecules. However, buffered solutions can be prepared in the laboratory, usually by combining a weak acid with the salt of that acid. The pH of these solutions can be calculated using the Henderson–Hasselbalch equation:

$$pH = pK_a + \log([\text{base}]/[\text{acid}])$$

Note that buffering is only effective at pH values equal to the $pK_a \pm 1$.

The buffering systems of the body do not excrete excess H^+ but temporarily remove them from free solution preventing excessive changes in pH. The effect of any H^+ produced by the body is neutralised largely by the $HCO_3^- - H_2CO_3$ buffer system, where:

$$H^+ + HCO_3^- \rightleftharpoons H_2CO_3$$

Therefore:

$$pH = pK_a + \log([HCO_3^-]/[H_2CO_3])$$

In plasma, H_2CO_3 breaks down to release CO_2 and water:

$$H_2CO_3 \rightleftharpoons CO_2 + H_2O$$

Since the $[H_2CO_3]$ is directly proportional to the partial pressure of CO_2 (PCO_2), it follows that:

$$[H_2CO_3] = PCO_2 \times 0.225$$

where 0.225 is the solubility constant of CO_2. The pK_a of the system is 6.1; hence, the Henderson–Hasselbalch equation can be rewritten as:

$$pH = 6.1 + \log([HCO_3^-]/PCO_2) \times 0.225)$$

It follows that the $[H^+]$ is directly proportional to the ratio $PCO_2 / [HCO_3^-]$. Hence, the $[H^+]$ in the blood varies as the $[HCO_3^-]$ and the value of PCO_2 change: an increase in H^+ occurs when there is an increase in PCO_2 or a decline in HCO_3^-, while a decrease in H^+ will occur when the PCO_2 decreases or HCO_3^- increases.

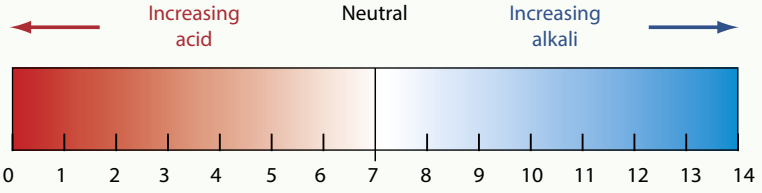


Figure 9.5 The pH scale.

Buffering by the $HCO_3^- - H_2CO_3$ system removes H^+ from the ECF but at the expense of HCO_3^-. The ECF contains relatively large amounts of HCO_3^-; its concentration is usually about 24 mmol/L. If, for any reason, the amount of H^+ produced increases, then the $[HCO_3^-]$ will decrease as the $HCO_3^- - H_2CO_3$ buffering system operates. Any excess H^+ must be excreted from the body by the kidneys (*Chapter 8*) to keep the ECF and tissues at an appropriate pH. In addition, the HCO_3^- used in buffering must be regenerated to return its concentration in the plasma to physiological values, otherwise the body will become depleted of HCO_3^- and buffering capacity.

The role of kidney glomeruli (*Figure 9.6*) in ultrafiltering the blood and the reabsorption of essential materials from the ultrafiltrate is described in *Chapter 8*. Two separate mechanisms operate in the kidneys to recover the HCO_3^- initially removed from the blood by filtration at the glomerulus and to regenerate that used in buffering: one is the HCO_3^- recovery system, while the second mechanism regenerates the HCO_3^- (*Figures 9.7* and *9.8*).

In health, nearly all the HCO_3^- is reabsorbed from the kidney tubule lumen. The operation of this biochemical system depends upon the tubule cells being polarised, that is, their luminal and basal surfaces differing in composition and permeability (*Figure 9.7*). Direct reabsorption of HCO_3^- from the renal tubule fluid cannot occur because the luminal surfaces of renal tubule cells are impermeable to HCO_3^-. However, the $[CO_2]$ within the tubule cells is maintained at a relatively high value and so carbonic anhydrase catalyses the formation of H_2CO_3. The acid dissociates to HCO_3^- and H^+. The continuous formation of HCO_3^- and H^+ within the tubule cells is promoted by their removal. The HCO_3^- is transported across the basal membrane of the cell into the interstitial fluid and then into

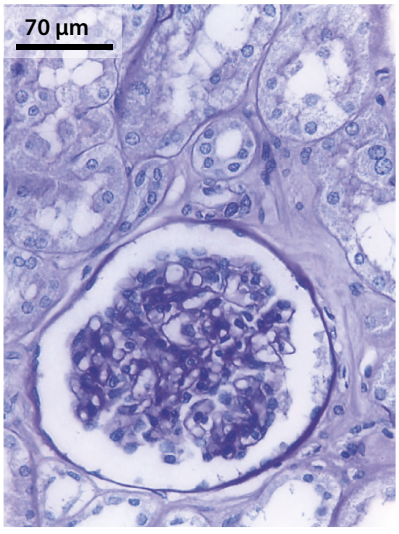


Figure 9.6 Photomicrograph of kidney tissue showing a glomerulus with sectioned kidney tubules.

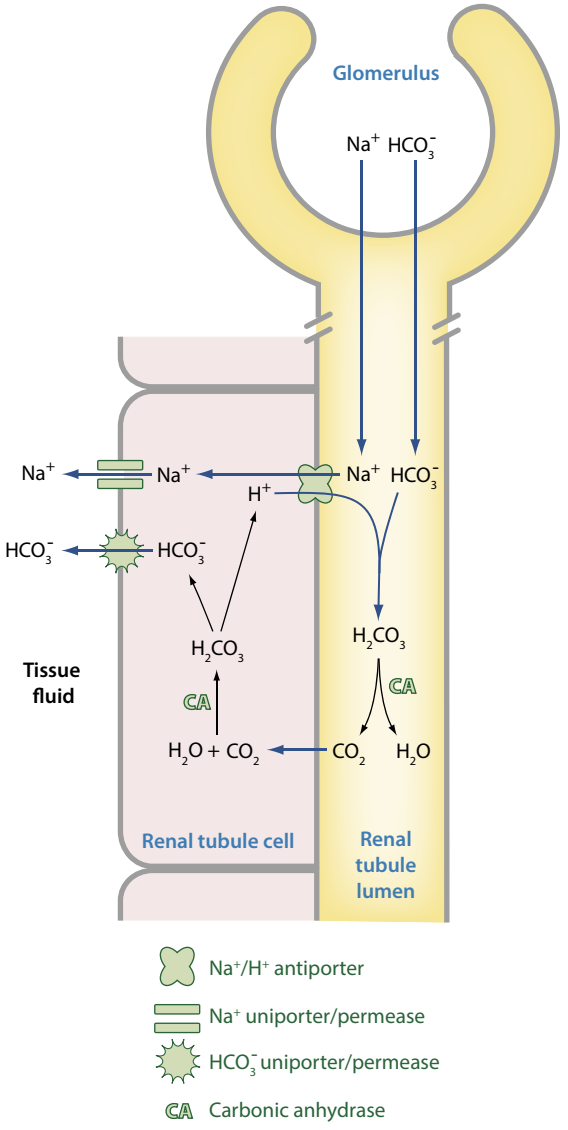


Figure 9.7 The indirect reabsorption of HCO_3^- by kidney tubule cells. Compare with *Figure 9.8*.

the capillaries. In contrast, H^+ is exchanged for Na^+ across the luminal membrane and enters the lumen of the kidney tubule. The sodium/bicarbonate cotransporter 1 (NBC 1, also referred to as SLC4A4), a protein present in the luminal cell plasma membrane, facilitates this exchange of ions. Within the lumen, the H^+ combines with HCO_3^- to form H_2CO_3. The acid breaks down spontaneously to CO_2 and H_2O in the proximal tubule, but carbonic anhydrase activity on the luminal surfaces of the cells speeds up the reaction in the distal tubule. The CO_2 can enter the tubule cell across its luminal membrane and so the HCO_3^- is recovered, indirectly, as CO_2.

This mechanism is effective in that the proximal tubule recovers approximately 80% of the filtered HCO_3^-. However, there has not been a net loss of H^+, and the HCO_3^- used in buffering has not been regenerated. The HCO_3^- is regenerated when H_2CO_3 is formed in the luminal cell as illustrated in *Figure 9.8*. Again, H^+ are exchanged for Na^+ and enter the lumen, where the H^+ react with phosphate (HPO_4^{2-}) and ammonia (NH_3) to give HPO_4^{2-} and NH_4^+, respectively. Ammonia, a significant urinary buffer, is produced by the deamination of glutamine in the renal tubule cells in a reaction catalysed by glutaminase (*Figure 9.9*). The ammonia formed can readily diffuse across cell membranes but NH_4^+ cannot enter the cells by passive reabsorption. Thus, the NH_4^+ and HPO_4^{2-} are excreted in the urine. For every H^+ excreted as NH_4^+ and HPO_4^{2-}, a single HCO_3^- is formed in the tubule cell and secreted across the basal surface to the interstitial fluid and then into the blood. Hence, the HCO_3^- concentration of the ECF is regenerated. The biosynthesis of glutaminase is induced in states of chronic acidosis (*Section 9.4*) allowing an increase in the production of ammonia and an increased excretion of H^+ as NH_4^+.

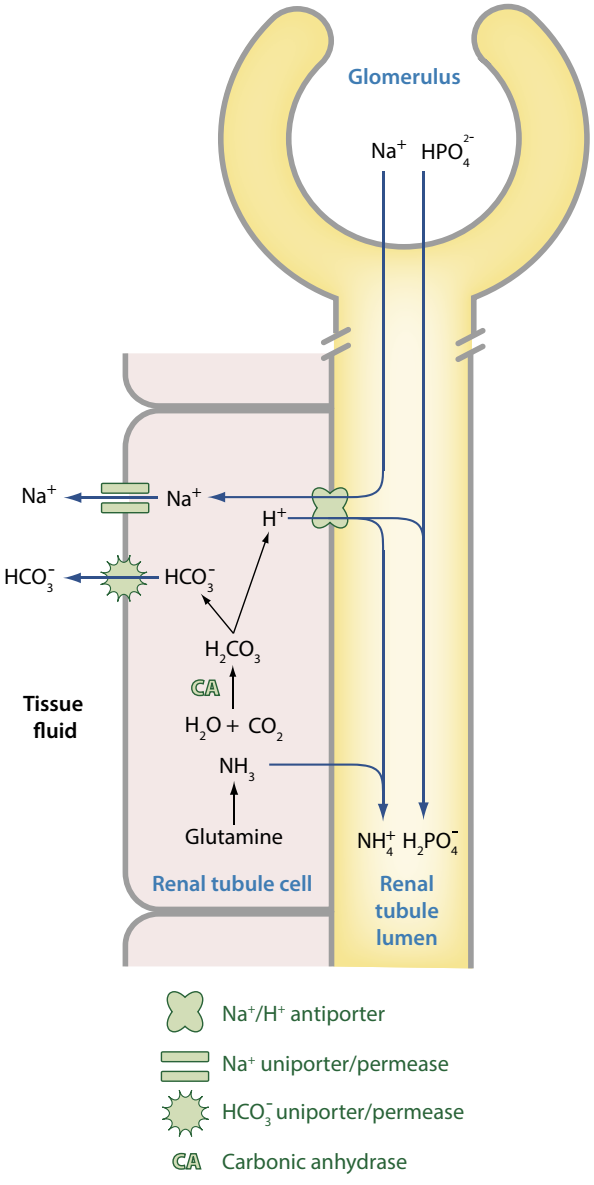


Figure 9.8 The regeneration of HCO_3^- by kidney tubule cells. Compare with *Figure 9.7*.

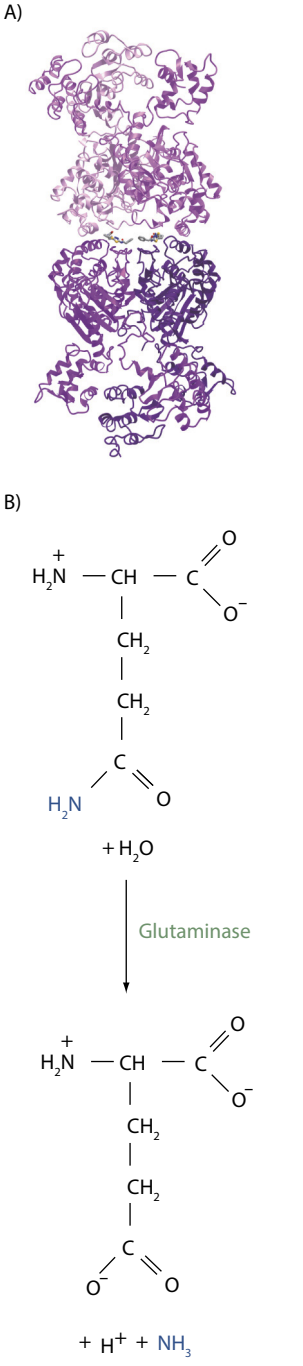


Figure 9.9 (A) Molecular model of glutaminase. PDF file 5UQE. (B) The formation of ammonia by the deamination of glutamine catalysed by glutaminase.

9.4 ACID–BASE DISORDERS

Disorders of acid–base balance are either *acidoses* or *alkaloses*. In an acidosis, H^+ accumulates in the blood and its plasma pH falls below its reference range. In an alkalosis, there is a depletion of H^+ and therefore the blood has a pH above its reference range. Acid–base disorders can be subdivided into two groups depending on their causes. If the abnormal pH occurs because of a metabolic or renal dysfunction, it is referred to as a **metabolic acid–base disorder**. When the abnormal pH is due to lung dysfunction, then it is a **respiratory acid–base disorder**. Physiological mechanisms that attempt to return the pH back to reference range values are referred to as **compensation**. Metabolic disorders cause a change in the $[HCO_3^-]$ in the blood, but respiratory disorders cause a change in its PCO_2 (*Table 9.1*).

Primary disorder	Effect	Compensatory response
Metabolic acidosis	Decreased $[HCO_3^-]$	Decreased PCO_2
Metabolic alkalosis	Increased $[HCO_3^-]$	Increased PCO_2
Respiratory acidosis	Increased PCO_2	Increased $[HCO_3^-]$
Respiratory alkalosis	Decreased PCO_2	Decreased $[HCO_3^-]$

Table 9.1 Types of acid–base disorders and their compensatory mechanisms.

In any acid–base disorder, the pH of the blood depends on the severity of the primary disturbance and the amount of compensation that has occurred. These can be caused by the production of ketone bodies during diabetic ketoacidosis (*Chapter 7*) or from the loss of HCO_3^- from a duodenal fistula. Respiratory acidosis is associated with an increased PCO_2, whereas respiratory alkalosis occurs when the PCO_2 is decreased. For example, an impairment of respiratory function can increase the PCO_2 in the blood while hyperventilation would decrease it.

Compensation of acid–base disorders occurs by two major mechanisms:

1. renal compensation
2. respiratory compensation

Renal compensation occurs when a respiratory disorder impairs lung function. The body attempts to adjust the pH of blood back to within its reference range by increasing the excretion of H^+ by the kidneys. Respiratory compensation is necessary when there is a metabolic acid–base disorder and involves changes in the ventilation of the lungs. Renal compensation is a relatively slow mechanism while respiratory compensation is much quicker to take effect. An acid–base disorder is said to be fully compensated if the compensatory mechanism returns the pH of the blood back to its reference range. However, compensation is usually only partial, and the pH remains outside the reference range.

Margin Note 9.2
Fistula

A fistula is an abnormal passage from a cavity or tube to another cavity or free surface. Compare to a hernia (see *Chapter 11*).

Metabolic acid–base disorders

Metabolic acid–base disorders lead to an accumulation or a loss of H^+ resulting in changes in the $[HCO_3^-]$ in the blood. The direct loss or gain of HCO_3^- will also cause a metabolic acid–base disorder. Thus, metabolic disorders are recognised by investigating the $[HCO_3^-]$ in the blood (see below and *Section 9.5*). Respiratory compensation occurs quickly, often within hours, and patients will show some change in blood PCO_2 because of hypo- or hyperventilation.

Metabolic acidosis

Metabolic acidosis may arise from an increase in the amount of H^+ formed or a decrease in the $[HCO_3^-]$. Diabetic ketoacidosis (*Chapter 7*), lactic acidosis, poisoning with for example ethylene glycol, methanol or salicylate (*Chapter 12*), or the condition inherited organic acid disorders (*Box 9.2*) all increase the production of H^+. In contrast, a decreased excretion of H^+ as in renal tubule acidosis, acute and chronic renal failure (*Chapter 8*), the use of inhibitors of carbonic anhydrase or a deficiency of mineralocorticoids, such as aldosterone (*Figure 8.6*), all increase blood H^+ content. Acid ingestion, as in acid poisoning, the excessive intake of amino acids by infusion, the direct loss of HCO_3^- by diarrhoea or a pancreatic fistula can all reduce the $[HCO_3^-]$ in the blood. The clinical effects of metabolic acidosis include hyperventilation, where the increased $[H^+]$ acts as a rapid and powerful stimulant of the respiratory centre, leading to deep sighing breathing called **Kussmaul respiration**. The patients may also present with neuromuscular irritability that can cause cardiac arrhythmias. Cardiac arrest is more likely in the presence of hyperkalaemia (*Chapters 8* and *14*). Eventually, metabolic acidosis can depress the activities of the central nervous system (CNS), and this can progress to coma and even death.

Patients with metabolic acidosis are managed by treating its underlying cause, which usually resolves the acid–base disorder. In severe cases, for example when the pH is less than 7.2, the patients may be administered HCO_3^- intravenously to correct the acidosis. Extreme cases of metabolic acidosis may require correction by dialysis (see *Figure 12.16*).

BOX 9.2 ORGANIC ACID DISORDERS

Organic acid disorders (OADs) or inherited organic acidurias are a class of inherited metabolic disorders (IMDs) that arise from defective/absent enzymes associated with intermediary metabolism. This results in an accumulation of organic acids, for example methylglutaconate and derivatives of butyric, crotonic, propionic and isovaleric acids (*Figure 9.10*), in tissues and their subsequent excretion in urine. The type of acid accumulated and excreted depends upon the nature of the metabolic defect. Individually, these are rare conditions, with even the most common having frequencies of 1 in 35 000–50 000. Clinical presentations vary with the type, but common signs and symptoms include infantile metabolic ketoacidosis, unexplained metabolic acidosis, failure to thrive and develop normally, vomiting, anaemias, seizures and other neurological abnormalities including coma, which can be fatal if untreated.

Analysis of organic acids in urine is essential for the diagnosis of organic acidurias. A diagnosis depends on determining the acid–base imbalance and analysing urine and sometimes blood samples to determine their organic acid and organic acid derivative contents. Most of the common organic acidurias can be effectively treated, but a diagnosis within the first 24–48 hours is vital, since rapid care can prevent both physical and mental developmental problems. Nutritional based therapies are necessary to support physical and mental development and prevent the accumulation of precursors of the toxic organic acids. Prognosis in many cases is excellent if diagnosis is made promptly, and the metabolic acidosis can be reversed especially if some residual enzyme activity is present.

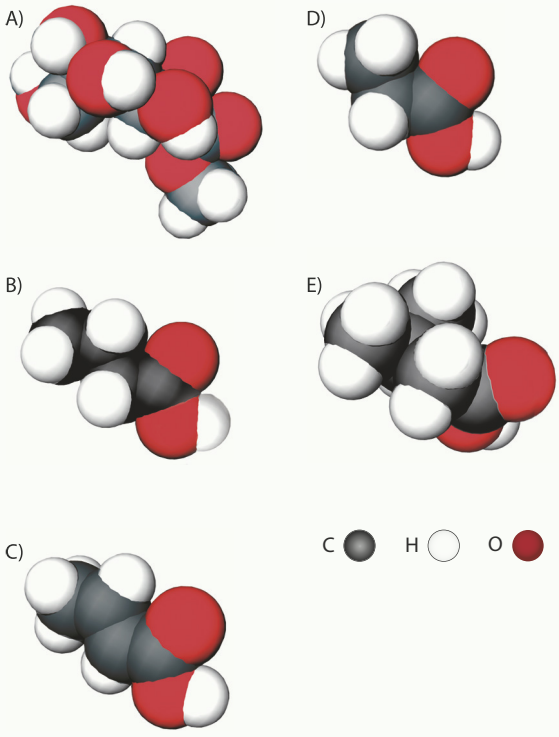


Figure 9.10 Computer drawn models of (A) the salt methylglutaconate and (B) butyric, (C) crotonic, (D) propionic and (E) isovaleric acids.

Metabolic alkalosis

Metabolic alkalosis may occur as a consequence of gastrointestinal loss of H^+ following vomiting and gastric aspiration or from excessive renal loss of H^+ in Conn and Cushing syndromes (*Chapter 7*). Some clinical treatments, such as the use of carbenoxolone, an anti-inflammatory drug used to treat ulcers, and thiazide diuretic drugs (*Figure 9.11*) that reduce blood pressure by promoting the production of urine and by K^+ depletion, can also result in this condition. Finally, the administration of alkali, including alkali ingestion, and inappropriate treatment for acidosis can also cause metabolic alkalosis.

The clinical effects of metabolic alkalosis include hypoventilation, which is a consequence of low blood $[H^+]$. It is often accompanied by mental confusion and eventually coma. Patients may also suffer from paraesthesia (numbness or tingling sensations). Other effects of metabolic alkalosis include tetany and muscle cramps that arise due to a decrease in the concentration of unbound Ca^{2+} in the plasma (*Chapter 8*).

The clinical management of a metabolic alkalosis is dependent on its underlying cause. For example, if vomiting is the cause, the use of an antiemetic, if possible, may help its correction. In patients suffering metabolic alkalosis due to treatment with certain diuretics, their dose may be reduced or stopped as appropriate.

Respiratory acid–base disorders

Respiratory acid–base disorders are primarily caused by a change in the partial pressure of arterial CO_2. These disorders are related to a defect in the rate of ventilation of lungs or in the exchange of gases across the alveolar membrane. The changes in PCO_2 alter the $[H_2CO_3]$ in the blood, which, in turn, dissociates to HCO_3^- and H^+ (see *Box 9.1*).

Some causes of respiratory acidosis are shown in *Table 9.2*. In general, obstruction of the airways by disease or inhibition of the respiratory centre in the brain by disease, trauma or drugs can cause respiratory acidosis.

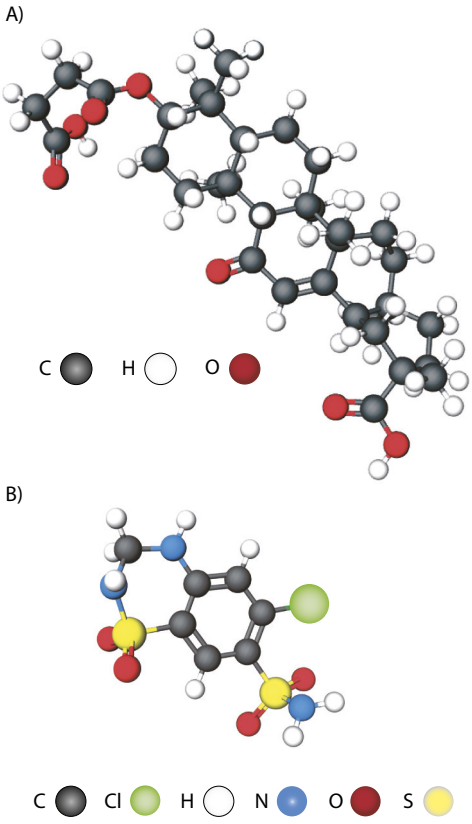


Figure 9.11 Computer drawn models of (A) the anti-inflammatory drug carbenoxolone and (B) the thiazide diuretic hydrochlorothiazide.

Conditions giving rise to respiratory acidosis	Examples
Chronic obstructive pulmonary disease (*Box 9.3*)	Bronchitis, emphysema
Obstruction of airways due to bronchospasms	Asthma
Inhibition of the respiratory centre	Anaesthetics, sedatives
Cerebral damage	Accidental trauma, stroke, tumours
Neuromuscular disease	Poliomyelitis, tetanus, Guillain–Barré syndrome
Pulmonary disease	Fibrosis, pneumonia, respiratory distress syndrome
Sleep apnoea	Obesity

Table 9.2 Some conditions giving rise to respiratory acidosis.

Respiratory acidosis

Respiratory acidosis may be acute or chronic. Acute conditions occur within minutes or hours. It is usually low PO_2 (**hypoxaemia**) that is more dangerous than high PCO_2 (**hypercapnia**). Further, renal compensation is slow, taking two or three days to become effective, so respiratory acidosis is usually uncompensated. Alveolar hypoventilation is usually the most common reason for acute respiratory acidosis. Hypoventilation increases the arterial PCO_2 and so the blood $[H^+]$ also rises quickly. The high PCO_2 and associated low PO_2 can cause coma and eventually death if untreated. Causes of acute respiratory acidosis include choking, bronchopneumonia and the acute exacerbation during an asthma attack (*Chapter 5*).

A chronic respiratory acidosis is, again, usually due to a decline in alveolar ventilation. However, this is normally a well-established condition and subject to maximum renal compensation. Long-standing conditions responsible for chronic respiratory disorders include chronic bronchitis and emphysema.

High PCO_2 is believed to be responsible for the clinical features of respiratory acidosis, such as peripheral vasodilatation and headaches. The acidosis can depress CNS activity leading to a coma. The treatment of respiratory acidosis is to improve alveolar ventilation, lowering the PCO_2 and increasing the PO_2. In chronic respiratory acidosis, it is usually not possible to treat the underlying cause, and treatment is aimed at maximising alveolar ventilation by using physiotherapy or bronchodilators.

Respiratory alkalosis

Respiratory alkalosis is less common than respiratory acidosis. However, it is often an acute condition due to hyperventilation. Often renal compensation does not occur.

The clinical effects of respiratory alkalosis include confusion, headaches, dizziness and coma. Respiratory alkalosis may be caused by hypoxia, increased respiration or pulmonary disease. Hypoxia is associated with high altitudes, severe anaemia and pulmonary disease. Increased respiration may result from the use of respiratory stimulants, such as salicylates, from the clinical condition primary hyperventilation syndrome, artificial hyperventilation and pulmonary diseases, such as pulmonary oedema and embolisms. The treatment of respiratory alkalosis is aimed at removing its underlying cause, as this usually resolves the acid–base disturbance. Patients can be given appropriate medication to reduce pain and anxiety and to correct the hyperventilation that underlies the respiratory alkalosis.

Mixed acid–base disorders

Sometimes patients may present with more than one acid–base disorder, known as a **mixed acid–base disorder**. These may present with any of the following:

- severe acidaemia, that is, a low blood pH
- a normal or near normal pH
- alkalaemia, that is, a high blood pH

Whatever the underlying cause, all mixed acid–base disorders are associated with abnormal levels of blood PCO_2 and HCO_3^-. For example, a patient with chronic bronchitis may also have renal failure. Both these disorders increase the $[H^+]$ in the blood. Chronic bronchitis leads to respiratory acidosis while the renal failure causes metabolic acidosis. This patient will therefore present with a mixed acid–base disorder with a high blood PCO_2 and $[H^+]$ but a low $[HCO_3^-]$. In some cases, however, the two disorders in a mixed acid–base disorder can be antagonistic, that is, have opposing effects on the $[H^+]$ in blood. In this case, the blood H^+ concentration may be near normal although the PCO_2 and $[HCO_3^-]$ will both be abnormal. For example, a patient with salicylate (aspirin) poisoning may have a metabolic acidosis together with a respiratory alkalosis. Patients may also present with metabolic and respiratory alkaloses. This could occur in a patient with congestive cardiac failure who is on diuretic therapy. The former will cause a respiratory alkalosis and the latter a metabolic alkalosis. Such individuals will usually have a high blood pH and increased HCO_3^- but the PCO_2 will be decreased.

BOX 9.3 CHRONIC OBSTRUCTIVE PULMONARY DISEASE ⏎

Chronic obstructive pulmonary disease (COPD) is a progressive, irreversible inflammatory lung disease characterised by diminished inspiratory and expiratory lung volumes (see *Chapter 5*) due to chronic bronchitis, emphysema and small airways disease. It is a common condition, which in 2021 caused over 3.5 million deaths, making it the third or fourth leading cause of death globally. In the UK, over a million cases of COPD have been diagnosed, although the number affected is thought to be much higher, and it is the second largest cause of emergency hospital admissions. Thus, after asthma (*Chapter 5*), COPD is the second most common lung disease. In the USA, COPD affects more than 15 million patients and has been reported as its fourth leading cause of death by the Centers for Disease Control and Prevention (CDC). Nearly 90% of COPD deaths occur in patients younger than 70 years. Research suggests that the incidence of COPD is increasing.

The major cause of COPD is tobacco smoking, with the risk of it developing increasing the more tobacco is smoked and the longer the duration of the habit. Atmospheric pollution is also a significant factor in its development. It has been estimated that 15–20% of cases result from long-term occupational inhalation of harmful fumes, vapours and microparticles. Less than 5% of COPD cases are associated with a deficiency in the glycoprotein α-1-antitrypsin (*Figure 9.12*) arising from mutations in the gene *SERPINA1*. This glycoprotein is produced in the liver and secreted into the blood, which transports it to the lungs. Here, it inhibits neutrophil elastase, a protease that can hydrolyse (digest) elastin of the alveolar walls. Elastase can be

Figure 9.12 Molecular model of α-1-antitrypsin. PDB file 1QLP.

secreted by neutrophils in pathological quantities following damage to lung tissues by exposure to tobacco smoke, some toxic chemicals and dust, and in inflammatory clinical conditions. A deficiency of α-1-antitrypsin can therefore exacerbate the effects of elastase damage leading to the development of an early onset form of COPD at 40–50 years of age.

Most COPD patients present with both emphysema and chronic bronchitis, although the seriousness of each condition is variable. Chronic bronchitis is a long-term inflammation of the linings of the trachea and bronchi, with an excessive production of mucus that can further restrict the flow of air. Emphysema is characterised by the degradation of elastic tissue of lung alveoli resulting in their collapse upon exhalation. This produces the typical signs and symptoms of COPD. These include a tight feeling in the chest, wheezing and difficulties in breathing, especially during physical activities; a chronic cough that may expel mucus ('phlegm'); and frequent respiratory infections. The inefficiency of respiratory gas exchange in COPD patients means it is a major cause of acid–base imbalances. The impaired excretion of CO_2 results in its excessive accumulation and the formation of H_2CO_3 in the tissues. The subsequent dissociation of the acid releases H^+, causing respiratory acidosis as described in the main text. However, this is not the only acid–base disorder associated with COPD, since the presence of comorbidities, such as renal failure, and the side effects of some drugs, for example the use of steroids to treat COPD patients, may induce metabolic alkalosis causing mixed acid–base disorders. In many cases, COPD increases the likelihood of other conditions developing, for example lung cancer (*Chapter 17*), heart arrythmia and failure (*Chapter 14*), and lung infections such as pneumonia (*Chapters 2* and *3*).

A diagnosis of COPD is based on taking a clinical history from the patient, especially in relation to tobacco smoking. Standard tests include spirometry measurements as described for asthma in *Chapter 5*. A chest X-ray will show several characteristic features, such as elongated, hyperinflated lungs with the presence of air pockets called bullae and a flattened diaphragm, and the heart may be narrowed (*Figure 9.13*).

There is no cure for COPD, and any lungs damage is permanent. Without treatment, the breathing problems associated with COPD usually become progressively worse. There may also be periods when they suddenly deteriorate, which are called flare-ups or exacerbations. These tend to become worse over time, although treatment can help keep the condition under control. Treatment is largely directed at the underlying cause(s). This is generally the cessation of tobacco smoking and increased physical activity to slow progression of the condition. Chronic obstructive pulmonary disease and asthma share a number of symptoms and, like asthma, the use of inhalers (see *Figure 5.26*) may ease breathing difficulties. Oxygen therapy and a specialised programme of exercises and education, referred to as pulmonary rehabilitation, may be necessary. Extreme cases may necessitate a lung transplant (*Chapter 6*).

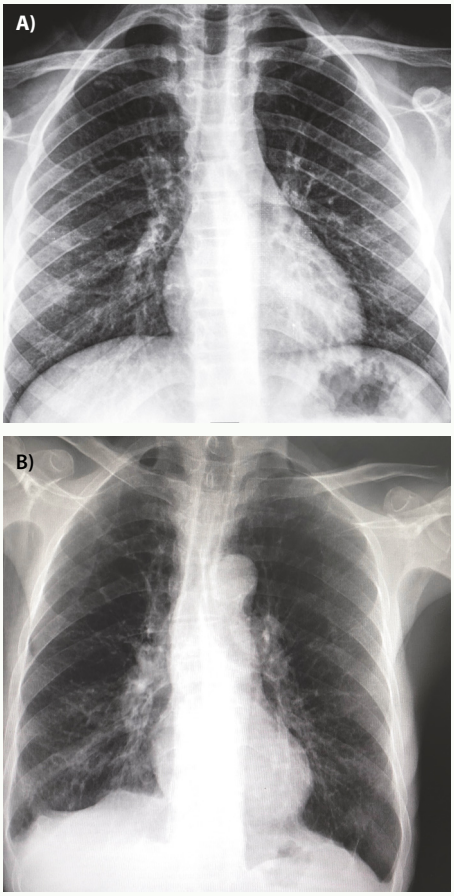


Figure 9.13 X-ray photographs of the chests of (A) a normal (healthy) person and (B) a patient suffering from chronic obstructive pulmonary disease (COPD). (A), Courtesy of Shutterstock.com/Miriam Gil Albert; (B), courtesy of Shutterstock.com/Chisanupong Chantapoon.

9.5 DIAGNOSING AND TREATING ACID–BASE DISORDERS

An investigation of an acid–base disorder involves three stages. The first involves identifying whether the patient has an acidosis or an alkalosis. The second stage is to determine whether the acid–base disorder is metabolic or respiratory in nature, while the third stage involves determining the degree of compensation. *Figure 9.14* outlines an algorithm that can investigate acid–base disorders. An arterial blood specimen is collected and its $[H^+]$, that is, the pH, and PCO_2 are measured. The blood must be collected from an artery into a syringe containing an anticoagulant, such as heparin, and transported to the hospital laboratory at 4°C. Care must be taken that air does not enter any blood samples during collection. The $[H^+]$ and PCO_2 levels are measured directly using an automated analyser that is also programmed to calculate the corresponding $[HCO_3^-]$ (*Figure 9.15*) according to the Henderson–Hasselbalch equation. An acidosis or alkalosis can then be identified from the pH of the blood. The values of PCO_2 and $[HCO_3^-]$ show whether the disorder is metabolic or respiratory in origin and indicates to what extent compensation is occurring.

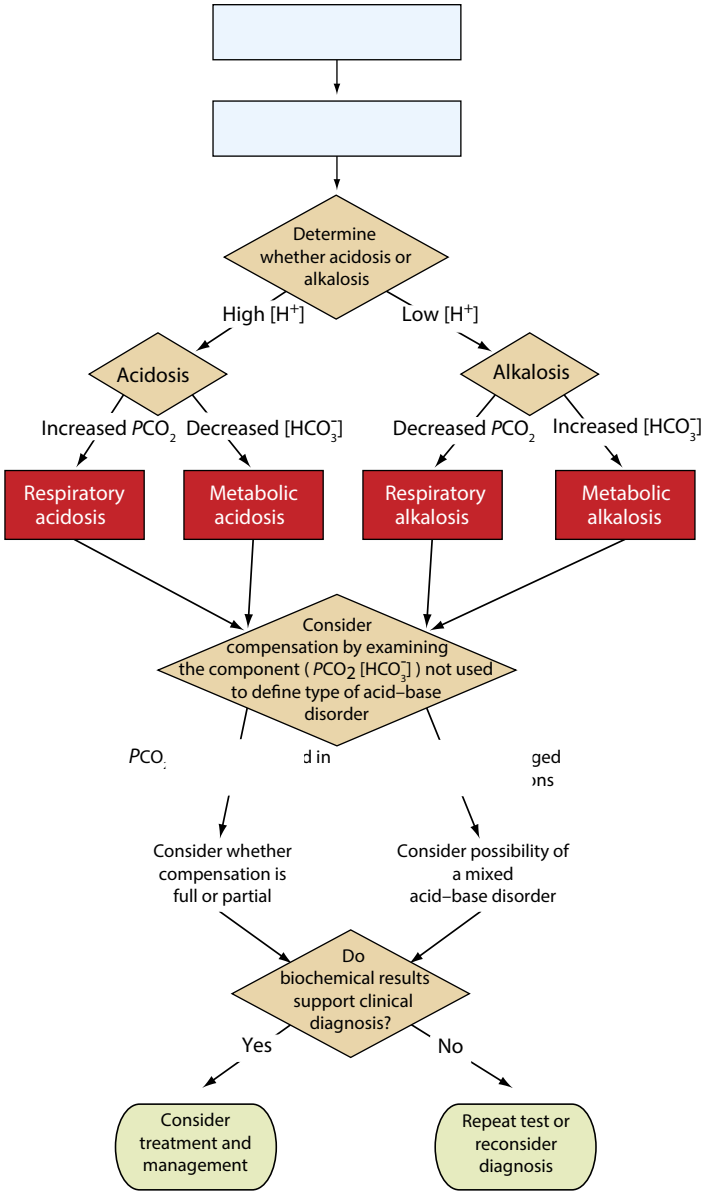


Figure 9.14 Outline of how disorders of acid–base balance are investigated.

Figure 9.15 A contemporary blood gas analyser that can determine 19 key parameters, including pH, PCO_2 and the corresponding $[HCO_3^-]$ using only a single 65 µL blood sample in 35 seconds. Courtesy of Radiometer Ltd, UK.

9.6 CASE STUDIES

CASE STUDY 9.1

John is a young man with a history of dyspepsia (severe indigestion) and excessive alcohol intake. He was admitted into his local hospital following 24 hours of vomiting. An arterial blood specimen was taken and analysed for a suspected acid–base disorder. The following results were obtained (reference ranges are indicated in parentheses):

$[H^+]$	28	(35–45 nmol/L)
PCO_2	7.2	(4.4–5.6 kPa)
$[HCO_3^-]$	48	(21–28 mmol/L)

Question

(a) Does John have an acid–base disorder and, if so, what type of acid–base disorder is present?

CASE STUDY 9.2

Tom was admitted to hospital following an acute attack of asthma. A blood specimen was collected and analysed, giving the following results (reference ranges are shown in parentheses):

$[H^+]$	24	(35–45 nmol/L)
PCO_2	2.5	(4.1–5.6 kPa)
$[HCO_3^-]$	22	(21–28 mmol/L)

Question

(a) Does Tom have an acid–base disorder and, if so, what type?

CASE STUDY 9.3

Terry, a 62-year-old factory worker, has suffered with vomiting and diarrhoea for two weeks. When examined by his doctor, he was dehydrated and his breathing was deep and noisy. An arterial blood specimen was taken and analysed for blood gases, giving the following results (reference ranges are shown in parentheses):

$[H^+]$	65	(35–45 nmol/L)
PCO_2	2.9	(4.1–5.6 kPa)
$[HCO_3^-]$	9	(21–28 mmol/L)

Questions

(a) Does Terry have an acid–base disorder?
(b) Is there any compensation?

CASE STUDY 9.4

Qiang is a heavy smoker and has suffered from chronic bronchitis for the last few months. He was admitted to hospital due to an excessive cough, breathlessness and wheezing. An arterial blood specimen was taken and analysed for a suspected acid–base disorder. The following results were obtained (reference ranges are shown in parentheses):

$[H^+]$	44	(35–45 nmol/L)
PCO_2	9.5	(4.1–5.6 kPa)
$[HCO_3^-]$	39	(21–28 mmol/L)

Questions

(a) Does Qiang have an acid–base disorder?
(b) If so, is there any compensation?

9.7 SUMMARY

The pH of the blood is maintained at 7.4 within normal limits by physiological mechanisms that remove the H^+ and CO_2 produced by metabolism. Excess CO_2 is transported by the blood to the lungs and excreted. In the blood, the dissolved CO_2 is converted to HCO_3^- in a reaction catalysed by carbonic anhydrase. Buffering systems such as haemoglobin and the $HCO_3^- - H_2CO_3$ system serve to maintain the physiological pH in health. Acid–base disorders result in the blood pH being lower (acidosis) or higher (alkalosis) than reference values. Such disorders may be metabolic or respiratory in nature and can be fatal if untreated, although there are physiological mechanisms for at least a partial compensation of the disorders. Most acid–base disorders are treated by addressing their underlying causes.

QUESTIONS

1. Approximately how many molecules of CO_2 are produced daily by oxidative metabolism? Avogadro's number is 6.02×10^{23}.
(a) 12×10^{24}.
(b) 1.2×10^{23}.
(c) 1.2×10^{26}.
(d) 1.2×10^{25}.
(e) 1.2×10^{24}.

2. **(a)** How much greater is the $[H^+]$ in a solution of pH 3 than one of pH 6?
(b) What is the pH of a buffer prepared from 100 mL of 0.1 mol/L ethanoic acid ('acetic acid') solution (pK_a of ethanoic acid is 4.76) and 75 mL of a 0.2 mol/L sodium ethanoate solution?

3. A blood analysis of a patient gives values for PCO_2 of 5.0 kPa (reference range: 4.4–5.6 kPa) and a pH of 7.56. The pK_a for the H_2CO_3 – HCO_3^- system is 6.10.
(a) Calculate the corresponding $[HCO_3^-]$.
(b) Is the patient in an acidosis or alkalosis?
(c) Does the patient's condition have a metabolic or respiratory basis?

4. The following blood gas results were obtained for a patient (reference ranges are shown in parentheses):

$[H^+]$	50	(35–45 nmol/L)
PCO_2	11.5	(4.4–5.6 kPa)
$[HCO_3^-]$	34	(21–28 mmol/L)

Which of the following conditions match most closely with these results?
(a) Diabetic ketoacidosis.
(b) Laboratory transcription error.
(c) Chronic obstructive pulmonary disease.
(d) Pyloric stenosis.
(e) None of the above.

5. Arrange the following two columns into their most appropriate pairings:

Protonated haemoglobin ('H^+Hb')	Blood pH > 7.4
Anaerobic glucose metabolism	Alveolar hypoventilation
6.1	Changed $[HCO_3^-]$
Na^+/H^+ antiporter	Metabolic acid–base disorder
NH_3	Kidney tubule cell
Alkalosis	Urine [butyric acid derivatives] increased
Metabolic acid–base disorder	$HCO_3^- - H_2CO_3$ buffer system
Respiratory compensation	Glutaminase
Organic acid disorders	General body tissues
Respiratory acidosis	Formation of ketone bodies

FURTHER READING

Allen K (2005) Four-step method of interpreting arterial blood gas analysis. *Nurs Times* **10**:42–45.

Ayers P *et al.* (2015) Acid-base disorders: learning the basics. *Nutr Clin Pract* **30**:14–20.

Berend K and Duits AJ (2019) The role of the clinical laboratory in diagnosing acid-base disorders. *Crit Rev Clin Lab Sci* **56**:147–169.

Bollmeier SG and Hartmann AP (2020) Management of chronic obstructive pulmonary disease: a review focusing on exacerbations. *Am J Health Syst Pharm* **77**:259–268.

Christenson SA *et al.* (2022) Chronic obstructive pulmonary disease. *Lancet* **399**:2227–2242.

Fall PJ (2000) A stepwise approach to acid-base disorders. Practical patient evaluation for metabolic acidosis and other conditions. *Postgrad Med* **107**:249–258.

Gomez H and Kellum JA (2015) Understanding acid base disorders. *Crit Care Clin* **31**:849–860.

Hamm LL *et al.* (2015) Acid-base homeostasis. *Clin J Am Soc Nephrol* **10**:2232–2242.

Tucker AM and Johnson TN (2022) Acid-base disorders: a primer for clinicians. *Nutr Clin Pract* **37**:980–989.

Wagner CA *et al.* (2019) Molecular pathophysiology of acid-base disorders. *Semin Nephrol* **39**:340–352.

Taylor & Francis
Taylor & Francis Group
http://taylorandfrancis.com

Diet and disease

10

OBJECTIVES

After studying this chapter you should be able to:

- define the term *nutrient*
- outline the types, structures and roles of nutrients
- list some common nutritional disorders
- explain the causes of some common nutritional disorders
- review some general methods used to investigate nutritional disorders
- discuss the treatments and management of some common nutritional disorders

10.1 INTRODUCTION

Nutrition is concerned with food and how the body uses it. Foods must be ingested, and most of them digested, before nutrients can be assimilated and used by the body (*Figure 10.1*). Carbohydrates, lipids and proteins are required in comparatively large amounts but other nutrients, such as vitamins, minerals and trace elements, although necessary, are required in much smaller quantities. Water is also necessary to support life.

Nutrients have a variety of metabolic roles and are needed for normal growth, development and the maintenance of health. The body can make some nutrients from others that must be supplied in the diet; the former are **nonessential** and the latter are **essential** nutrients. Specific daily nutritional requirements are determined by a number of factors and, indeed, vary throughout the life of any individual. The major factors concerned include the age, sex, physical activity and general wellbeing of the person concerned and, in the case of females, menstruation, pregnancy and lactation. Good nutrition is essential for health and the prevention of clinical conditions and diseases. Nutritional disorders may arise from a deficiency or, in some cases, an excess of some component of the diet.

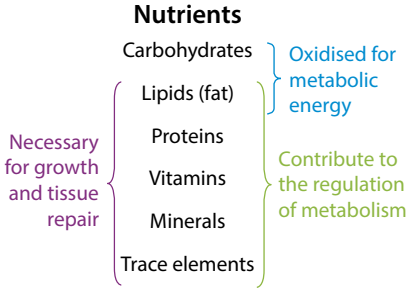


Figure 10.1 Overview of the general roles of nutrients.

10.2 DIET AND NUTRITION

Food is obviously needed to sustain life, but a balanced diet is vital for good health. General evidence for good nutrition is a well-developed body with an ideal weight (*Section 10.3*), healthy skin including hair and nails, a satisfactory muscle tone, and physical and mental alertness. The food eaten by an individual constitutes the diet, and it should supply appropriate amounts and proportions of necessary nutrients. Macronutrients, such as carbohydrates, lipids and proteins, are required in amounts that are large compared with the micronutrients, that is, vitamins, minerals and trace elements, that are needed in much smaller quantities.

Nutrients are needed to sustain a number of activities within the body if good health is to be maintained. They supply the raw materials required for growth and maintenance of tissues, including substances that help regulate metabolism, for example the cofactors and coenzymes necessary for enzyme activities. They also provide the energy to drive metabolism and maintain homeothermy. Nutrients vary in their contributions to these roles and there is also overlap between their contributions to each (see *Figure 10.1*).

DOI: 10.1201/9781003296836-10

Nutrients and energy

Diets are often described in terms of their energy content, despite the requirement to supply materials for functions other than just energy provision. It is possible for a diet to supply an adequate amount of energy but be deficient in some essential raw materials, such as vitamins and minerals needed by the individual. A **balanced diet** is one where the food (*Figure 10.2*) supplies adequate energy distributed appropriately between carbohydrates, lipids and proteins and contains the necessary amounts of vitamins, minerals, trace elements, water and nondigestible fibre.

Nutritional energy is usually measured in joules or kilojoules. The number of joules required by any one person needs to be matched to their energy output (*Table 10.1*). The term **recommended daily amount** (RDA) was originally used by the UK Department of Health in 1979 to define the amounts of certain nutrients needed by different groups within the population. Unfortunately, RDAs were often used inappropriately to assess the dietary needs of individuals. Accordingly, in 1991, they were replaced with **dietary reference values** (DRVs), which are guidance values as to the amounts of energy and nutrients that should be ingested, rather than exact recommendations. Dietary reference values indicate the amount of energy or quantity of an individual nutrient required by a group of a certain age range and sex for good health. A DRV is given as a daily intake. However, in practice, DRVs must be determined from the energy and nutrient intakes averaged over several days of eating because food intake and appetite vary from day to day. Groups for which DRVs have been set include seven groups of infants of both sexes from birth to age 10 years and four groups each for males and females between the ages of 11 and 50-plus years, with extra groups for women who are pregnant or lactating. The DRV applies to healthy people, since patients who are unwell for whatever reason may have differing nutritional requirements.

Carbohydrates
bread
cereals
pasta
peas
potatoes
rice

Fats
butter
cheese
cooking oils
cream
lard
margarine
milk
nuts
oily fish
suet

Proteins
cheese
eggs
fish
meat
milk
potatoes
pulses

Figure 10.2 A selection of foods that would supply the carbohydrate, fat and protein nutrients necessary for a balanced diet. Some foods, of course, contain more than one of these macronutrients and all will supply vitamins, minerals and trace elements.

Group	Approximate daily energy requirements per desirable body weight (kJ/kg)
Babies up to 1 year old	500
Children 1–10 years old	335–420
Male children 11–15 years old	270
Female children 11–15 years old	145
Sedentary males and most females	117
Physically active males and females	125
Males and females over 55 years old	117
Pregnant females (first trimester)	117–134
Pregnant females (second and third trimesters)	150–159
Lactating females	150–159

Table 10.1 Energy required to maintain body functions.

Dietary terms related to DRVs include:

- **estimated average requirement** (EAR)
- **reference nutrient intake** (RNI)
- **lower reference nutrient intake** (LRNI)
- **safe intake**

The EAR is the mean amount of energy or nutrient needed by a population, while the RNI is an amount of a nutrient that is greater than the dietary needs of 97.5% of that group. The LRNI is the amount of a nutrient sufficient to meet the requirements of 2.5% of a population. Note that most people will need more than the LRNI. *Figure 10.3* shows how the EAR, RNI and LRNI are related. Finally, the safe intake is the amount judged to be sufficient for the needs of most people and anything below this level could be undesirable. This arbitrary estimate is given when there is simply insufficient evidence to establish reliable values for the EAR, RNI or LRNI.

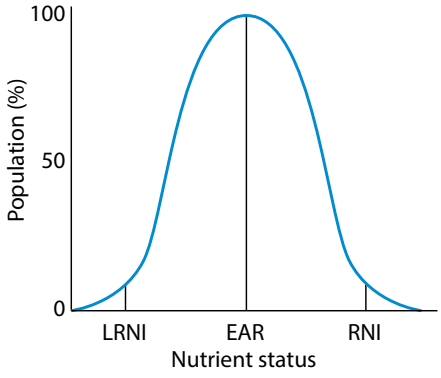


Figure 10.3 The relationships between estimated average requirement (EAR), reference nutrient intake (RNI) and lower reference nutrient intake (LRNI).

Carbohydrates

Carbohydrates have the general formula $(CH_2O)_n$. The main dietary carbohydrates are mono- and disaccharide sugars, for example fructose, glucose (see *Margin Note 10.1*), lactose and sucrose, and polysaccharides, mainly starch, and these are usually the major suppliers of energy. In 2022, the World Health Organization (WHO) recommended that most dietary carbohydrate should be in the form of starch, with free sugars constituting less than 10% of the total energy intake. However, dietary fibre or roughage largely consists of cellulose, a polysaccharide that cannot be digested or absorbed by humans, although it may be metabolised by bacteria in the large intestine (*Chapter 11*). Foods rich in fibre include wholemeal cereals and flour, leaf and root vegetables, pulses, nuts and fruits. Dietary fibre stimulates peristalsis, protects against constipation and is known to reduce blood cholesterol, glucose concentrations and the incidence of colorectal cancers (*Chapter 17*). Communities that consume high-fibre diets have relatively low incidences of these cancers.

Lipids

Dietary lipids are fats and oils, both of which contain fatty acids that are concentrated sources of energy. Fats are also needed as carriers of the fat-soluble vitamins A, D, E and K (see later). Dietary fatty foods consist largely of triacylglycerols together with small amounts of phospholipids and cholesterol. The WHO recommendations of 2022 included that total fat intake should not exceed 30%, with saturated fats and trans fats (see below) forming less than 10% and 1%, respectively, of dietary energy.

Essential fatty acids

Fatty acids may be saturated or unsaturated, that is, they lack or contain double bonds, respectively. The only difference between fats and oils is in their melting points, which are related to their degree of unsaturation and the carbon chain lengths of their constituent fatty acids (see *Margin Note 10.2*): oils are liquid at room temperature, indicating a higher degree of unsaturation than fats, which are solid at this temperature.

Humans lack the enzymes necessary to synthesise certain unsaturated fatty acids, termed **essential fatty acids** (EFAs), which are necessary for health. These include linoleic, linolenic and arachidonic acids, which are precursors of many biologically active and clinically relevant molecules, such as eicosanoids (prostaglandins, thromboxanes, prostacyclins and leukotrienes). These molecules mediate a wide range of physiological activities affecting, for example, inflammatory responses, blood pressure and clotting, reproductive activities and the sleep–wake cycle. Their actions resemble those of hormones but are local, affecting only cells near their sites of production, and they act at much lower concentrations, although via second messengers (*Chapter 7*).

Rather than the strict chemical convention of numbering carbon atoms in fatty acids from the carboxyl group, the double bonds in fatty acids are often numbered from the terminal or ω carbon atom (*Figure 10.5*). The term 'omega fatty acid' has entered everyday English. Humans, like all mammals, lack the enzymes needed to form fatty acids with double bonds between C-9 and the terminal methyl group. This metabolic limitation gives rise to three families of fatty acids with their first double bond occurring at positions ω-3, ω-6 and ω-9 and which cannot be metabolically interconverted.

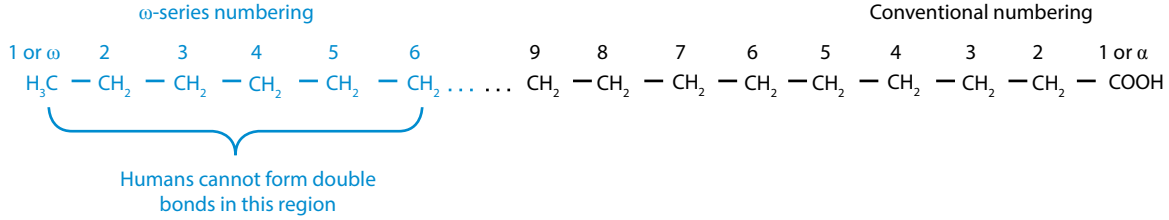

ω-series numbering
Conventional numbering
1 or ω 2 3 4 5 6 9 8 7 6 5 4 3 2 1 or α
H3C — CH2 — CH2 — CH2 — CH2 — CH2 . . . . . CH2 — CH2 — CH2 — CH2 — CH2 — CH2 — CH2 — CH2 — COOH
Humans cannot form double bonds in this region

Figure 10.5 The numbering of carbon atoms in fatty acids.

Since they cannot be synthesised by the body, ω-3 and ω-6 fatty acids (*Figure 10.6*) must be obtained in the diet or synthesised from precursor ω acids in foods, for example the two forms of linolenic acid: an ω-3 type called α linolenic acid (ALA) and the ω-6 form, γ linolenic acid (GLA). Seeds and vegetable oils are excellent dietary sources of ω-6 EFAs and the body is able to convert linoleic acid to GLA and arachidonic acid. However, many

Margin Note 10.1 Glucose and the glycaemic index

Glucose (*Figure 10.4*) is considered a universal fuel, given all organisms can metabolise it to obtain energy.

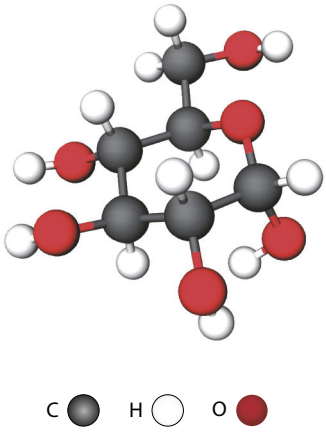


Figure 10.4 Computer drawn structure of α glucose.

The glycaemic index (GI) ranks foods from 1 to 100, with glucose arbitrarily rated at about 100, based on how quickly a food eaten on its own affects the concentration of glucose in the blood. Foods with high GI values are digested and the resulting nutrients absorbed relatively quickly compared with those of low values. The former are generally high in carbohydrate, such as sugar and sugary foods and drinks, white bread and rice, and potatoes. The latter includes leafy vegetables, pulses, nuts and wholegrain foods, for example porridge oats, which are all digested slower, leading to a more gradual increase in blood sugar concentration.

Dieticians generally regard foods with lower GI values as healthier options, and rating foods in this manner can help in planning diets for certain clinical groups, such as diabetics.

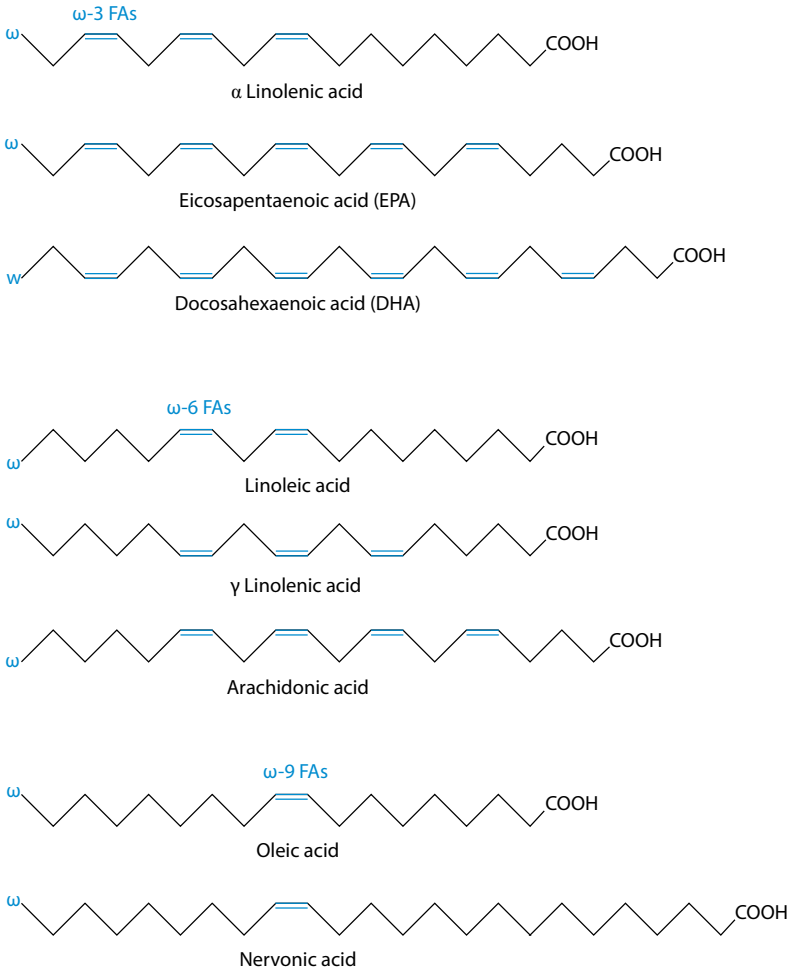


Figure 10.6 Examples of ω-numbered unsaturated fatty acids (FAs).

vegetable oils contain only low amounts of ALA, since this is normally present in the chloroplasts of plants and may only be a significant dietary constituent if green leafy vegetables are consumed. Body tissues can use ALA to synthesise eicosapentaenoic acid (EPA) and docosahexaenoic acid (DHA). Rich sources of ω-3 EFAs, EPA and DHA, are the oils from the muscles and skins of a number of cold deepwater fishes, such as herring, mackerel, salmon, sardines and tuna (*Figure 10.7*). Cod, which live in similar environments, store fat in the liver: hence cod liver oil is also an excellent source of EPA and DHA.

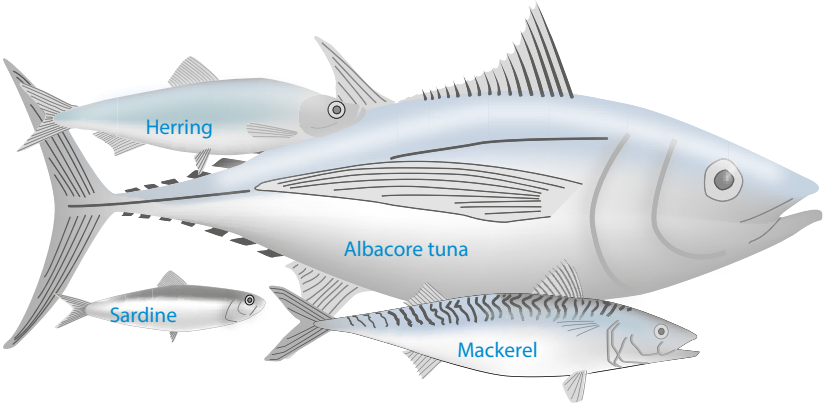


Figure 10.7 A selection of fish rich in essential fatty acids.

Margin Note 10.2
Weak acids (-ics) and their salts (-ates)

Fatty acids, such as arachidonic acid and some vitamins, for example vitamins B_5 (*Figure 10.8*) and B_9 (pantothenic and folic acids, respectively) are weak acids in their native states, that is, they are *-ics*. However, in the buffered, near-neutral pH of most biological fluids, they are ionised and occur as salts, that is, they are *-ates* (see also *Margin Note 8.3*).

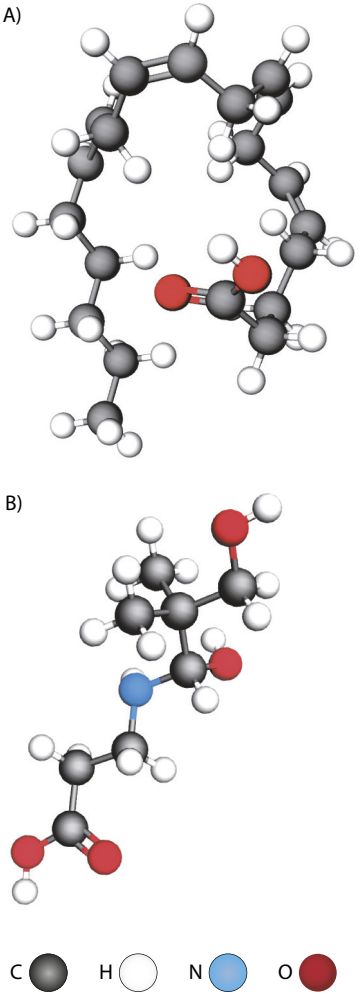


Figure 10.8 Computer drawn structures of (A) arachidonic acid and (B) vitamin B_5 (pantothenic acid).

A dietary deficiency of EFAs leads to dry scaly skin subject to erythema, poor healing of wounds, hair loss and a failure to thrive in infants. The first associations between ω fatty acids and health came from studying Greenland Inuit and inhabitants of fishing villages in Japan and people of Okinawa. These groups have much lower incidences of diseases, such as coronary heart disease (CHD; *Chapter 14*), rheumatoid arthritis (*Chapters 5* and *18*) and diabetes mellitus (*Chapter 7*), than their European or USA counterparts, even though their traditional diets are high in fat from eating seals and fish. However, these organisms are rich in ω-3 fatty acids that have been shown subsequently to provide significant health benefits. The metabolites of ω-3 fatty acids decrease platelet function, reduce the risk from sudden death caused by cardiac arrhythmias and slow the progress of atherosclerosis. Modest decreases in blood pressure also occur with high intakes of ω-3 fatty acids. Some studies have indicated that ω-3 fatty acids inhibit the biosynthesis of very low-density lipoprotein (VLDL) and triacylglycerols (TAGs) and so decrease their concentration in plasma, particularly in patients suffering from hypertriglyceridaemia (*Chapters 13* and *14*). Furthermore, they have been associated with reducing morning stiffness and the number of tender joints in patients with rheumatoid arthritis. There is, however, no established recommended daily intake for ω-3s, although it has been suggested that 0.5–2% of the total daily energy intake should be in the form of EFAs. Human adults are thought to require up to 10 g of linolenic acid daily.

The eicosanoids formed from ω-6 fatty acids, although produced by many of the same enzymes as those from ω-3 acids, have different physiological functions and effects. For example, the ω-3 group is antithrombic and anti-inflammatory, but both processes are promoted by ω-6 acid-derived metabolites. Thus, a balancing of the intakes of ω-3 and ω-6 fatty acids in the diet is thought to be of importance. The diet typical of the developed world is far richer in ω-6 than ω-3 fatty acids; it has been suggested that ratios of ω-3:ω-6 of 2:1 or 1:1 are a healthier balance. The American Heart Association and the UK Food Standards Agency recommend that people without a history of CHD should consume two servings of fish weekly and those with known CHD should eat one serving of fish daily. More than these amounts are not recommended because, unfortunately, significant amounts of environmental contaminants, such as methylmercury, polychlorinated biphenyls and dioxins, are concentrated in some species of fish. For the same reason, young children and women who may become pregnant, who are pregnant or who are breastfeeding should avoid eating excessive amounts of fish and shellfish. Omega-3 fatty acids exert a dose-related effect on bleeding time, and it has been suggested that care be taken to avoid excessive intake, particularly when combined with anticoagulant medications, such as warfarin or heparin (*Chapter 13*). In general, however, their ingestion is well tolerated, although side effects can include a fishy aftertaste and gastrointestinal tract (GIT) disturbances, for example nausea, bloating and belching, all of which appear to be dose dependent.

Proteins

Dietary proteins are necessary to supply the amino acids needed for growth and the general repair and maintenance of tissues. A daily intake of about 65 g and 50 g of dietary protein is required in males and females, respectively, which provides about 10–15% of the total energy in a balanced diet, although only about 5% of body energy comes from their catabolism under normal circumstances. Protein must be obtained from a variety of sources to supply all essential amino acids. Humans are unable to biosynthesise 9 of the 20 amino acids found in proteins that have codons in the genetic code and these are therefore essential dietary constituents (*Table 10.2*). However, the nonessential amino acids can be formed if the supply of the essential ones is adequate. The 'quality' of dietary protein is important, and protein intake needs to be varied, particularly as some plant proteins lack one or more of the essential amino acids.

Water

Water is vital for life because it helps stabilise the structures of large molecules, such as proteins and starches; functions as a medium for most biochemical reactions; acts as a solvent for electrolytes, glucose, vitamins, minerals and numerous other small molecules; and transports nutrients to, and waste products away from, cells as well as around the whole body. The GIT (*Chapter 11*) secretes some 7–9 L of water daily that aids the digestion and absorption of nutrients. To maintain these functions, loss of body water should match intake to ensure water balance (*Chapter 8*). It is recommended that 2–3 L of water should be consumed daily, of which about 60% should be liquid water and the rest obtained from seemingly solid foods. The homeostasis of water is described in *Chapter 8*.

Essential	Nonessential
H	A
I	R[a]
L	N
K	K
M[b]	C
F	E
T	Z
W	G
V	P
	S
	Y[c]

[a] R can be synthesised by tissues but not in sufficient amounts to support growth in children.
[b] M is required in substantial dietary quantities to form C if this is not ingested in adequate amounts.
[c] Increased quantities of E are needed to form Y if this is not ingested in adequate quantities.

Table 10.2 Essential and nonessential amino acids.

Vitamins

Vitamins are organic substances that the body cannot biosynthesise or can make only from chemically closely related compounds. They are needed in the diet but only in relatively small amounts (*Table 10.3*).

Vitamin or related compounds	Recommended intake (mg/day)	Sources
B_1 (thiamine)	1.0 in males 0.8 in females	Yeast, eggs, milk, cheese and cereals
B_2 (riboflavin)	1.3 in males 1.1 in females	Liver, milk, cereals, mushrooms and eggs
B_3 (niacin or nicotinic acid, nicotinamide)	17 in males 13 in females	Meat, fish, pulses, cereals and synthesised endogenously from tryptophan
B_5 (pantothenic acid)	3–7 is normally provided in the diet and appears more than adequate	Meat, offal, eggs, green vegetables and whole grains
B_6 (pyridoxine, pyridoxal, pyridoxamine)	1.4 in males 1.2 in females	Vegetables, fruit, cereals, eggs, milk and nuts
B_7 (biotin, H)	0.01–0.2 in adults	Eggs, pulses, liver, yeast, milk, dried fruit and can be synthesised by intestinal bacteria in the GIT
B_9 (folic acid)	0.2 in adults 0.4 in pregnancy	Liver, spinach, cabbage, peas, bananas and oranges
B_{12} (cobalamin)	0.0015 in adults	Liver, eggs and milk
C (ascorbic acid)	40 in adults	Green vegetables and citrus fruits
A (retinol, retinal, ß carotene)	0.7 in males 0.6 in females	Liver, fish oils, eggs, cheese and can be synthesised from carotene in the diet
D (cholecalciferol, ergocalciferol)	0.01 in pregnancy	Liver and dairy products and endogenous synthesis in the skin if there is adequate exposure to sunlight
E (tocopherol, tocotrienols)	4 in males 3 in females	Vegetable oils, nuts and wheat germ
K (phylloquinone [K_1], menaquinones [K_2])	0.001 in adults	Leafy vegetables, cheese or synthesised by colon bacteria

GIT, gastrointestinal tract.

Table 10.3 Recommended intakes of vitamins and some of their sources.

Vitamins (*Figure 10.10*) have specific biochemical roles and are essential for normal metabolism, growth and good health.

Vitamins may be classified as water soluble, such as the B group and vitamin C, and fat soluble, for example vitamins A, D, E and K. Some of the B vitamins and, in general, each fat-soluble vitamin comprise a group of closely related compounds called **vitamers**. In such cases, the name of the vitamin is used as a collective descriptor.

Water-soluble vitamins

Vitamin B_1 (thiamine) is an essential component of the coenzyme or prosthetic group thiamine pyrophosphate (TPP; *Figure 10.11*). This coenzyme is necessary for the activity of some enzymes, for example transketolase, in the pentose phosphate pathway for glucose oxidation and the oxidative decarboxylations catalysed by pyruvate and 2-oxoglutarate dehydrogenases in carbohydrate metabolism. Thiamine is therefore necessary for the metabolic formation of ATP, the major energy carrier in metabolism, and NADPH used in reductive reactions in biosyntheses. Thiamine pyrophosphate is also known to function in nerve conductance.

Margin Note 10.3
Selenocysteine

The amino acid selenocysteine (*Figure 10.9*) is a component of a number of selenoproteins, including the enzymes glutathione peroxidase, tetraiodothyronine 5′-deiodinase, thioredoxin reductase, formate dehydrogenase, glycine reductase and a number of hydrogenases. Selenocysteine is specified by the codon UGA, which would normally function as a termination codon in protein synthesis. However, in the presence of a selenocysteine insertion sequence element (SECIS) in the mRNA, UGA specifies selenocysteine. Characteristic nucleotide sequences and base-pairing patterns in the SECISs form secondary structures in the mRNA that lead to the change in codon specificity.

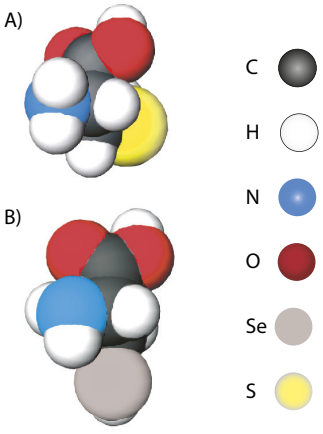


Figure 10.9 Computer drawn models of (A) cysteine and (B) selenocysteine.

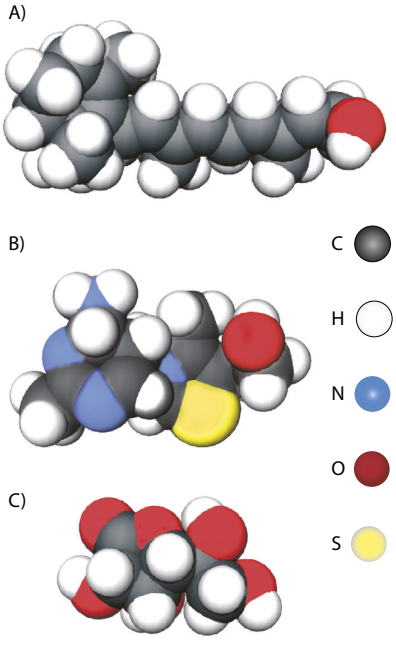


Figure 10.10 Computer drawn models of vitamins (A) A, (B) B_1 and (C) C.

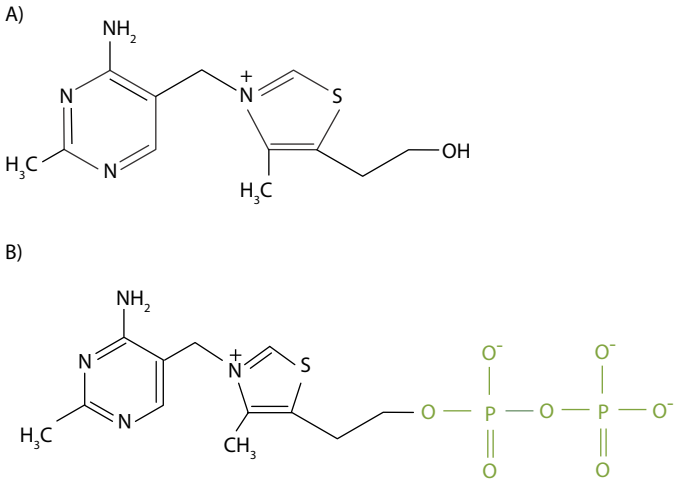


Figure 10.11 The structures of (A) vitamin B_1 (thiamine) and (B) thiamine pyrophosphate (TPP).

Vitamin B_2 (riboflavin) is also required to form coenzymes or prosthetic groups, in this case flavin mononucleotide and flavin dinucleotide (FMN and FAD; *Figure 10.12*), which function as electron carriers in flavoproteins. Flavoproteins are extremely numerous. They include renal L-amino acid oxidase, NADH reductase, 2-hydroxyacid oxidase (FMN), D- and L-amino acid oxidases, succinate dehydrogenase and glutathione reductases. Hence, riboflavin is essential to many oxidation and reduction pathways in, for example, the TCA cycle, electron transport and the β oxidation of fatty acids in mitochondria.

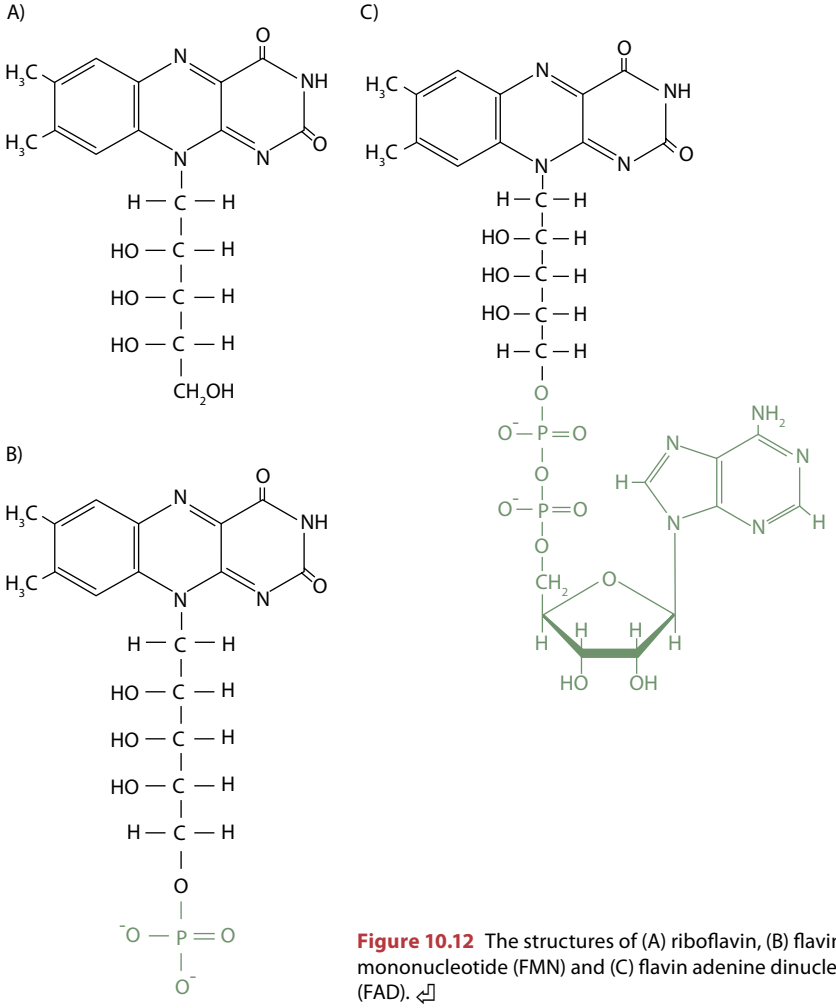


Figure 10.12 The structures of (A) riboflavin, (B) flavin mononucleotide (FMN) and (C) flavin adenine dinucleotide (FAD).

Vitamin B_3 (niacin or nicotinic acid) and its vitamer nicotinamide can be regarded as honorary vitamins, since limited amounts of it can be synthesised from tryptophan via kynurenine. Niacin is required to form the coenzymes NAD^+ and $NADP^+$, which are electron and hydrogen carriers (*Figure 10.13*). The former is crucial in electron transport associated with oxidative phosphorylation and ATP formation; the reduced form of NADPH is essential for the biosynthesis of, for example, sugars, lipids, amino acids and nucleotides.

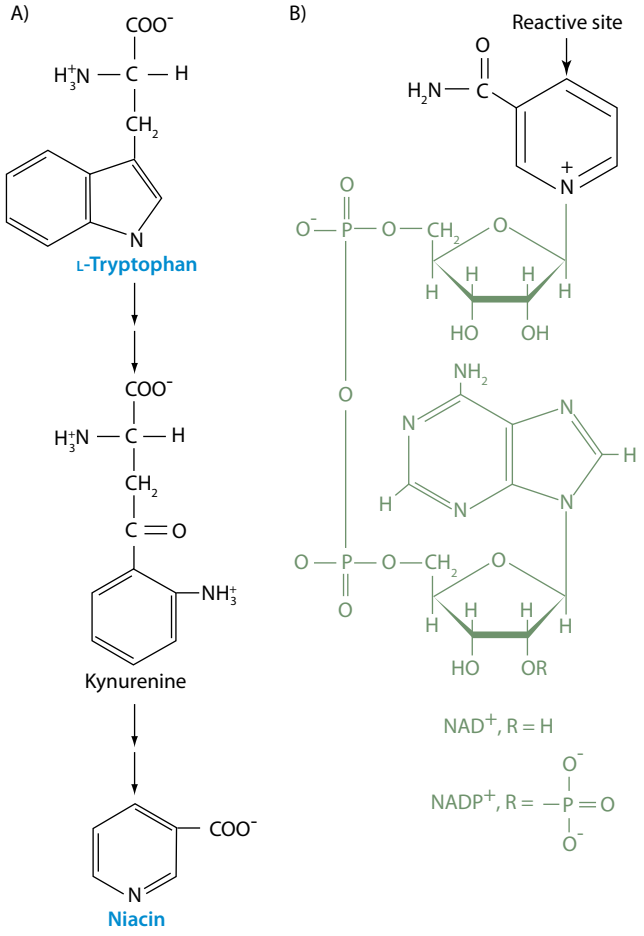


Figure 10.13 (A) The conversion of tryptophan to vitamin B_3 (niacin or nicotinate) and (B) the structures of NAD^+ and $NADP^+$.

Vitamin B_5 (pantothenic acid) is an essential part of coenzyme A (*Figure 10.14*), which is the major carrier of metabolically active acyl (fatty acid residues) groups in metabolism. Thus, it is essential to many of the reactions involved in the oxidation and in the biosynthesis of lipids, including steroid hormones, some neurotransmitters and haemoglobin.

Vitamin B_6 activity is displayed by its vitamers, pyridoxine, pyridoxal and pyridoxamine (*Figure 10.15*), forming pyridoxal 5-phosphate, which is a cofactor for a number of enzymes. These include glutamate decarboxylase, which catalyses the formation of γ aminobutyrate, a neurotransmitter of the central nervous system, and enzymes that catalyse transamination and decarboxylation reactions of amino acids. Vitamin B_6 is therefore essential for the biosynthesis of nonessential amino acids and in the catabolism of amino acids. Pyridoxal 5-phosphate is also a cofactor for glycogen phosphorylase of liver and muscle tissues and helps regulate the actions of steroid hormones (*Chapter 7*) by participating in the dissociation of hormone–receptor complexes from DNA.

Vitamin B_7 (H or biotin; *Figure 10.16*) is required to form the prosthetic group that carries CO_2 in a number of enzymes. These include acetyl CoA carboxylase and pyruvate decarboxylase, which are key enzymes of fatty acid synthesis and gluconeogenesis, that is, the production of glucose from noncarbohydrate precursors.

Vitamins B_9 (M or folic acid) and B_{12} (cobalamin) have a close functional relationship and, to some extent at least, depend on each other for activation. Their structures are

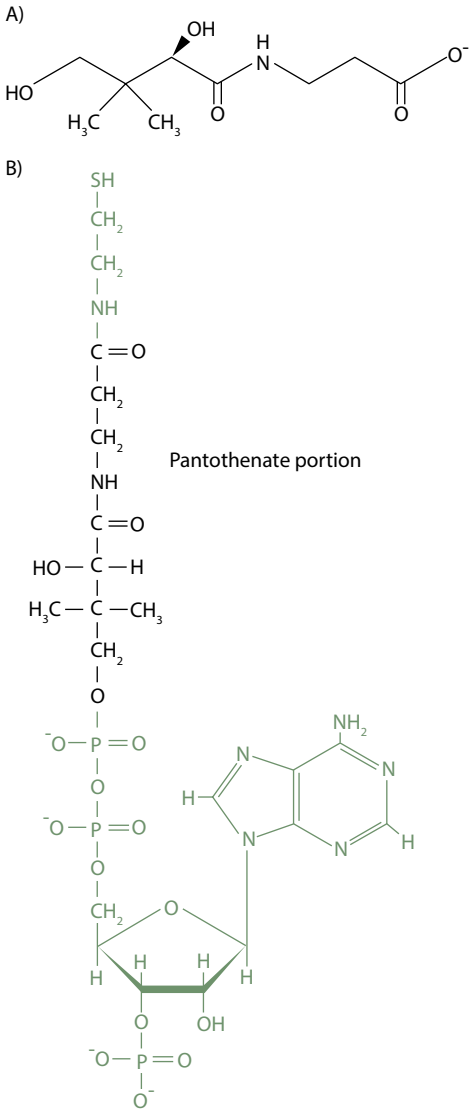
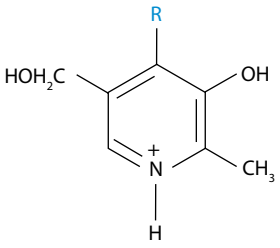


Figure 10.14 The structures of (A) vitamin B_5 (pantothenate) and (B) coenzyme A.



Pyridoxamine, R = CH_2NH_2
Pyridoxal, R = CHO
Pyridoxine, R = CH_2OH

Figure 10.15 Vitamers of vitamin B_6.

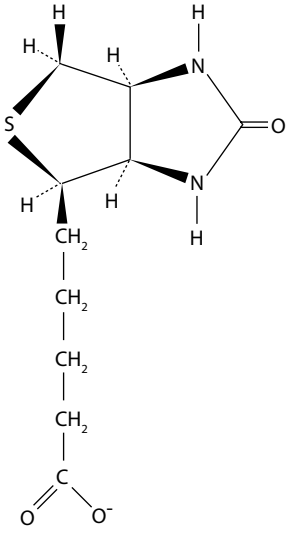


Figure 10.16 Vitamin B_7 (vitamin H or biotin).

illustrated in *Figure 10.17*. Note the unusual organometallic bond involving the cobalt ion in the cobalamin molecule. Unlike biotin that carries CO_2, folate and cobalamin carry organic one carbon compounds, for example methyl (CH_3–), methylene (–CH_2–), methenyl (–CH=), formyl (–CHO), formate ($-COO^-$) and formino (–CHNH). These compounds are generally toxic but during metabolism they are bound to the folate-cobalamin derived carriers, which allows them to be converted to different oxidation states and used in a variety of different reactions in a nontoxic manner. These reactions are necessary for the catabolism of some amino acids, the formation of a number of proteins and the synthesis of purine and pyrimidine bases and therefore nucleotides and nucleic acids.

Vitamin C (ascorbic acid; *Figure 10.18*) is essential to reduce metal ions in a number of enzymes following catalysis. Prolyl and lysyl hydroxylases contain Fe^{2+}, which is oxidised to Fe^{3+} during hydroxylation reactions involved in cross-linking the collagen molecules that add strength to connective tissues. Ascorbate reduces their iron back to the Fe^{2+} state, regenerating an active enzyme. Similarly, the Cu^{2+} in enzymes involved in synthesis of catecholamine hormones (*Chapter 7*) is returned to the Cu^+ state following the oxidation of the copper during catalysis. The antioxidant properties of ascorbic acid, in association with vitamin E, help protect lipids in the cell membranes and blood lipoproteins from oxidative damage. It also enhances the absorption of iron and regulates the absorption of copper from the GIT.

Fat-soluble vitamins

Vitamin A activity is shown by the vitamers retinol, retinaldehyde or retinal and retinoic acid. Retinol can be metabolically converted to retinaldehyde, which, in turn can

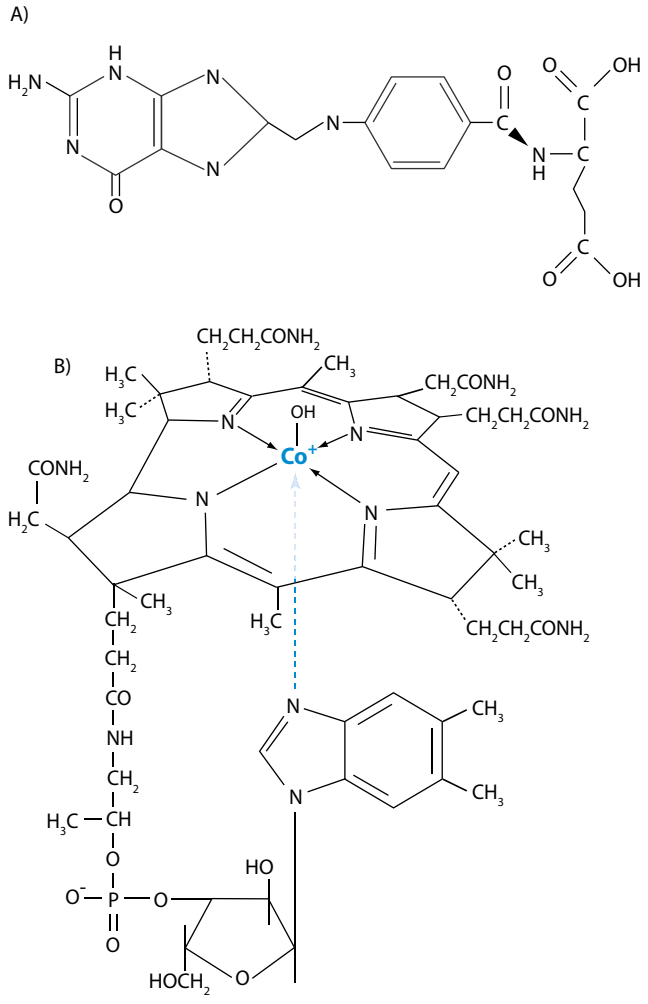


Figure 10.17 Vitamins (A) B_9 (folic acid, M) and (B) B_{12} (cobalamin).

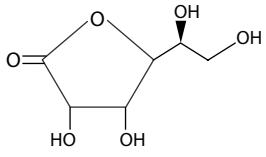


Figure 10.18 Vitamin C (ascorbic acid).

be oxidised to retinoic acid. In addition, the provitamin A carotenoids, for example β carotene (*Figure 10.19*), can be converted to active forms in the liver. Retinoate helps regulate the proliferation and development of cells in a tissue-specific manner. Its actions resemble those of steroid hormones (*Chapter 7*) in that it binds to nuclear receptors and activates specific genes. Vitamin A is associated with the development of epithelial cells,

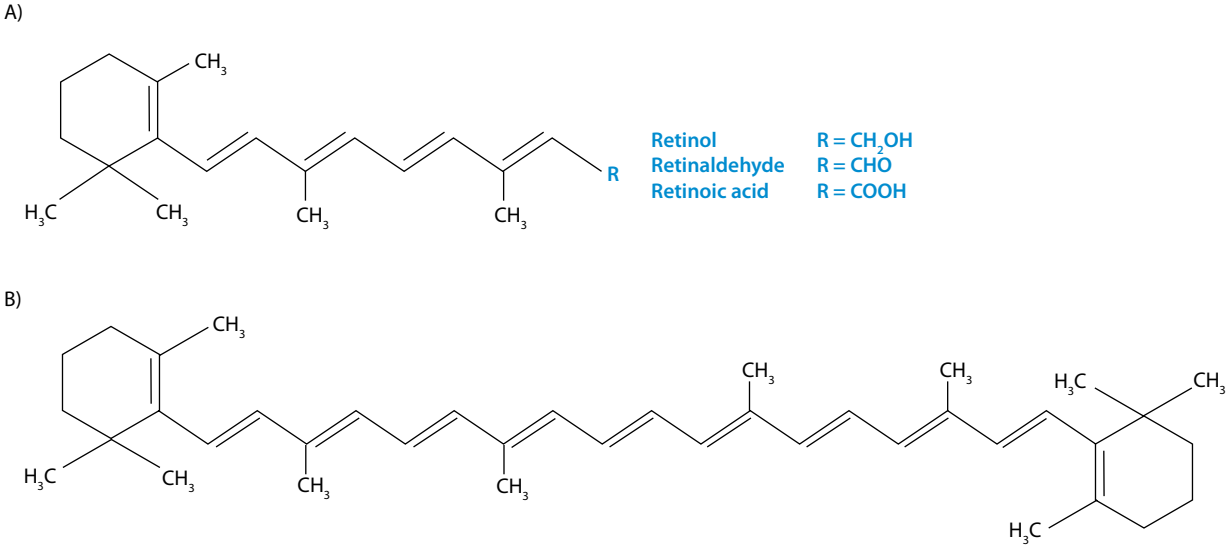


Figure 10.19 The structures of (A) vitamers of vitamin A and (B) β carotene.

such as skin and the mucosal membranes that cover external and internal surfaces of the body, and have numerous functions, for example as structural barriers that prevent microorganisms from entering the body (*Chapters 4* and *11*). Retinaldehyde is necessary for vision and functions as the prosthetic group (visual pigments) of opsin proteins in light-sensitive retinal cells. Vitamin A is a weak antioxidant that can protect against free radical damage.

Vitamin D_3 (cholecalciferol) is the usual dietary form of the vitamin, although this is not strictly a vitamin, since it can be formed by ultraviolet irradiation of the skin from 7-dehydrocholesterol (*Figure 10.20*). Foods are sometimes fortified with the synthetic ergocalciferol, which has the same biological activity as cholecalciferol. Enzyme-catalysed hydroxylations yield the active metabolites 1α,25-dihydroxycholecalciferol and calcitriol, respectively. Vitamin D functions mainly in the homeostasis of calcium as described in *Chapter 8*.

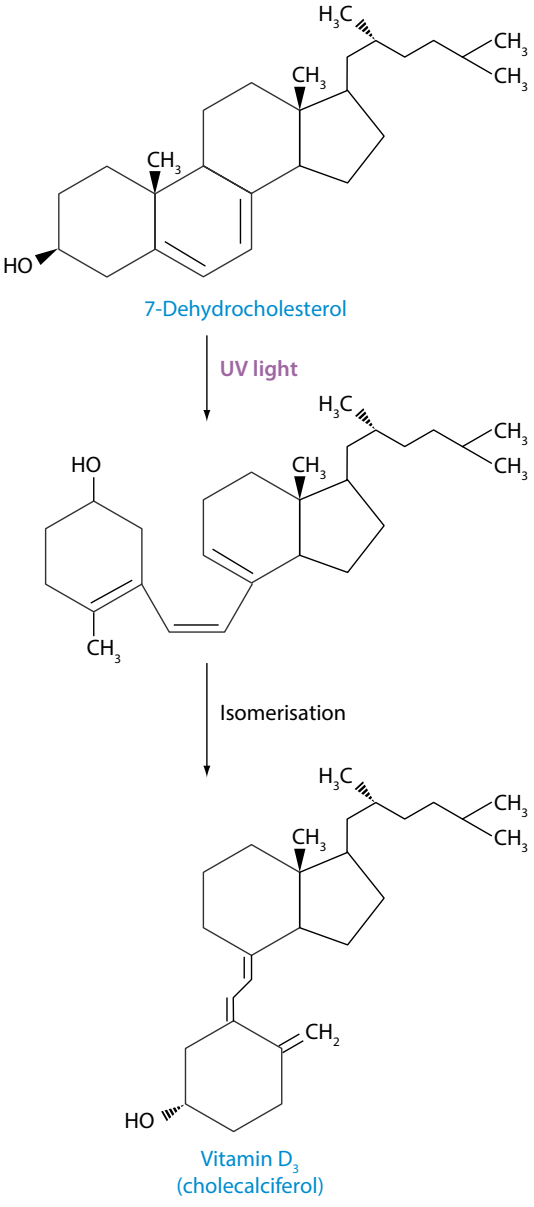


Figure 10.20 The conversion of 7-dehydrocholesterol to vitamin D_3 (cholecalciferol). UV, ultraviolet.

Vitamin E is generally used to describe the tocopherols and tocotrienols that comprise a number of vitamers of differing biological potencies, of which the most active is α-tocopherol (*Figure 10.21*). Like vitamins A and D, E also has roles in signal transduction and regulating gene expression. For example, it binds to pregnane X receptors in hepatocytes leading to an increased expression of enzymes concerned with vitamin E metabolism. A major function generally ascribed to vitamin E is to protect cellular

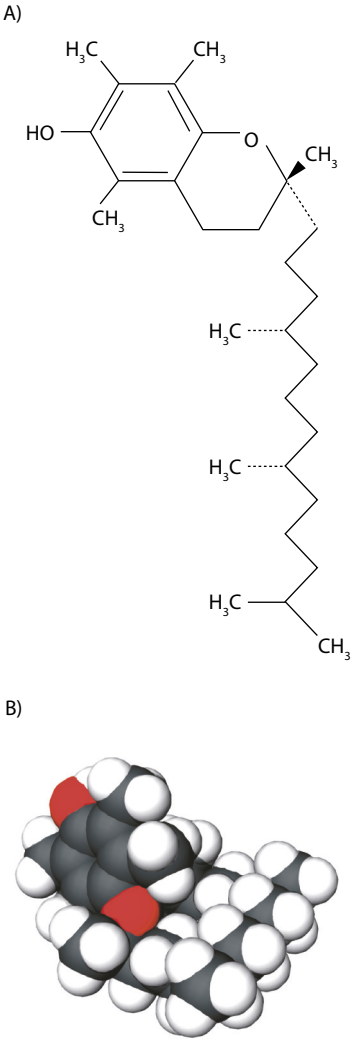


Figure 10.21 (A) The structure of and (B) a computer drawn molecular model of vitamin E (α-tocopherol).

membranes against free radicals (described in *Chapters 12* and *18*) and preventing the oxidation of plasma lipoproteins, especially low-density lipoproteins (*Chapter 14*). The corresponding reduction of the vitamin produces a relatively unreactive, and therefore less damaging, tocopheroxyl radical. This is also relatively long-lived and so persists sufficiently long to be reoxidised back to the active form by vitamin C or glutathione peroxidases (*Figure 10.22*).

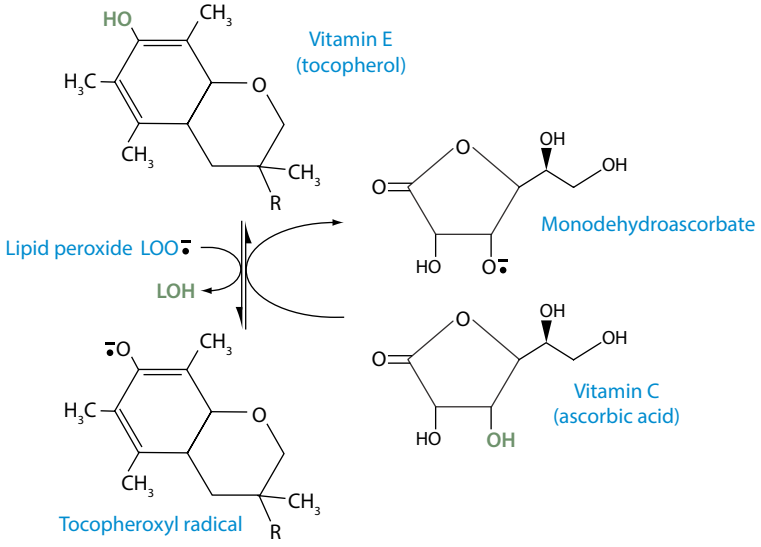


Figure 10.22 The roles of vitamins E and C in removing free radicals, such as lipid peroxides (LOO*) that could damage membrane lipids.

Vitamin K comprises a number of vitamers, including phylloquinone (K_1) and the menaquinones (K_2). The first is the normal dietary source, while the menaquinones are compounds with similar structures with carbon chains of varying lengths that are synthesised by GIT bacteria. Phylloquinone (*Figure 10.23*) is metabolically necessary for the post-translational modification conversion of glutamate residues to γ carboxyglutamates, which is necessary for the biosynthesis of some blood-clotting factors (*Chapter 13*) and some proteins of the bone matrix. Some menaquinones function as vitamin K but their significance is unclear; they may partially satisfy the human requirements for vitamin K and, indeed, they form a substantial proportion of the vitamin K stored in the liver.

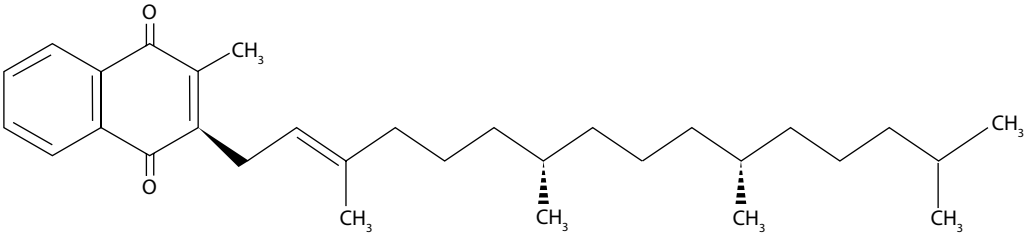


Figure 10.23 Vitamin K_1 (phylloquinone).

Minerals and trace elements

Minerals and trace elements are inorganic dietary substances required to maintain good health. Minerals include calcium, magnesium, sodium, potassium, chloride, phosphate and sulphate. They are present in the body in amounts larger than 5 g and some are required in dietary quantities of more than 100 mg per day and are provided by a variety of foods (*Table 10.4*).

Both minerals and trace elements have diverse functions. They play roles in promoting growth and are required constituents of body tissues such as bone, teeth, skin, hair and nails, and as cofactors in some enzymes and other proteins.

Minerals

Sodium, potassium and chloride are required to maintain the electrolyte and osmotic composition of intra- and extracellular fluids, to generate electrochemical gradients across

Mineral	Recommended daily intake (mg/day)	Sources
Calcium	700 in adults	Milk, cheese, vegetables, nuts, fish
Magnesium	300 in males 270 in females	Vegetables, nuts, bread, fish, meat
Sodium and chloride	<6000 in adults	Milk, cheese, salt
Potassium	3500 in adults	Bananas, vegetables, nuts, fish, pulses, poultry
Phosphate	550 in adults	Meat, fish, poultry, rice, oats
Sulphate	Balanced diet adequate	Milk, meat, legumes, eggs

Table 10.4 Recommended intakes and sources of minerals.

biological membranes and for nerve conductance and muscle contraction. Calcium, an essential component of bone and teeth, is also required for muscle contraction and is a second messenger for some hormones and neurotransmitters (*Chapter 7*). Magnesium is a cofactor for many enzymes, especially those involved in energy metabolism that utilise nucleotides, and in nucleic acid biosynthesis. Like calcium, phosphate is needed for bone and teeth (*Chapter 11*) formation and is also a component of nucleotides, nucleic acids and phospholipids. It activates a number of enzymes, especially some involved in energy metabolism, and is an important renal buffer. Sulphate is required for the synthesis of cartilage and other components of the extracellular matrix.

Trace elements

Trace elements are required in much smaller quantities than minerals but, like them, are supplied by a variety of foods (*Table 10.5*). Trace elements include chromium, cobalt, copper, fluoride, iodide, iron, manganese, molybdenum, selenium and zinc. They are present in the body at concentrations less than 100 parts per million and are required in milligrams or even micrograms per day, and a number are toxic in excess. Trace elements have specific and diverse functions.

Trace element	Recommended intake (mg/day)	Sources
Chromium	0.025 in adults	Meat, cheese, water, whole grains, lentils, spices
Cobalt	0.0015 in adults	Beef, eggs, offal, fish, milk products, nuts, vegetables, cereals
Copper	1.2 in adults	Offal, fish, nuts
Fluoride	Diet generally adequate	Present in drinking water in many areas of the world
Iodide	0.14 in adults	Fish, milk, eggs, cereals, iodised salt
Iron	8.7 in males 14.8 in females	Liver, meat, eggs, beans, nuts, dried fruit, brown rice, green vegetables, fortified cereals
Manganese	Diet generally adequate	Bananas, egg yolk, bread, nuts, cereals, vegetables, tea
Molybdenum	Diet generally adequate	Whole grains, leguminous vegetables, meat
Selenium	0.075 in males 0.06 in females	Fish, meat, offal, Brazil nuts, eggs
Zinc	5.5 to 9.5 in males 4.0 to 7.0 in females	Meat, offal, eggs, fish, cereals, milk, cheese, spinach, beans

Table 10.5 Recommended intakes and sources of trace elements.

Chromium helps maintain blood glucose concentration by acting as a cofactor for insulin activity. The role of cobalt as a component of vitamin B_{12} has already been mentioned (see above). Copper is also an essential cofactor for a number of enzymes, including those involved in collagen and elastin biosyntheses and some redox proteins. Copper is also required for iron absorption and metabolism and haemoglobin synthesis (*see Box 10.1*).

BOX 10.1 INHERITED DISORDERS OF COPPER METABOLISM

Menkes or kinky hair disease is a neurodegenerative disease of impaired copper transport that results in extreme tissue copper deficiency, which was first described in 1962. It has an incidence of 1 in 50 000–250 000. Menkes disease is a sex-linked inherited disorder and only affects male children (*Chapter 15*). The condition arises from mutations in *ATP7A* found on the X chromosome (Xq21.1), which encodes an integral membrane copper-transporting adenosine triphosphatase (ATPase) that transfers Cu^{2+} across the membrane. Transport of dietary copper from intestinal cells is impaired in Menkes disease, leading to low plasma copper concentrations with poor incorporation of copper into cuproproteins. Copper accumulates in excessive amounts in the liver but is deficient in most other tissues and impairs the functions of a number of essential enzymes, including tyrosinase, monoamine oxidase, cytochrome *c* oxidase, lysyl oxidase and ascorbic acid oxidase. Low tyrosinase activity leads to depigmentation of hair, while deficient monoamine oxidase results in kinky hair. Impaired lysyl oxidase function leads to defects in elastin and collagen, resulting in vascular weaknesses, and cytochrome *c* oxidase malfunction results in hypothermia. Defective ascorbate oxidase activity leads to skeletal demineralisation and osteoporosis.

Babies born with classic Menkes disease appear normal at birth and symptoms typically begin about three months later. The disease is characterised by a failure to thrive, psychomotor deterioration, seizures, hypothermia and strikingly peculiar hair, which is characteristically kinky, stubby, tangled, sparse or steely and easily broken and is often white, ivory or grey coloured. Brain problems, such as blood clots at the base of the brain (subdural haematoma) and/or rupture or thrombosis of arteries in the brain, may occur. Recurrent respiratory and urinary tract infections are common and weakened bones (osteoporosis) may result in fractures.

A diagnosis of Menkes disease in neonates could be based on high concentrations of copper levels in the placenta. However, this would be unusual, given a general absence of signs and symptoms at this age. After six weeks, low serum levels of copper and the copper binding protein ceruloplasmin (see *Figure 11.32*) are indicative of the condition. A microscopic examination of hair samples shows characteristic Menkes abnormalities, and a skin biopsy can be used to assess copper metabolism. There is no effective treatment for Menkes disease and prognosis is poor. Milder forms of the disease respond to intramuscular injections of copper, but the severe form does not show much change. Other treatments are focused on relieving the symptoms. Most patients die within the first decade of life; however, survival to the late 20s has been reported. Genetic screening of the individual's family can identify carriers and provide guidance for counselling on recurrence risks.

Wilson disease is characterised by an excessive accumulation of copper, although patients have normal or even low concentrations of copper in their blood and an increased urinary excretion. It was first described in 1912. Wilson disease, although more common than Menkes disease, is still rare, with an incidence of 1 in 30 000. It is an autosomal recessive condition caused by mutations in *ATP7B* located at 13q14.3, which, like *ATP7A*, encodes a plasma membrane Cu^{2+}-transporting ATPase. Indeed, the two proteins are over 75% homologous in primary structure. However, *ATP7A* is predominantly expressed in the placenta, gastrointestinal tract (GIT) and endothelial cells of the blood–brain barrier (see *Figure 3.8*), while *ATP7B* expression is mainly hepatic and renal. In contrast with *ATP7A*, mutations lead to an accumulation of copper in the liver and brain and also damage to other tissues.

The first symptoms shown by about 50% of patients with Wilson disease are swelling and tenderness of the liver and sometimes fever; this resembles more common disorders, such as viral hepatitis. Abnormal levels of circulating liver enzyme activities can indicate serious liver damage, which can progress to cirrhosis. Other symptoms may include jaundice, abdominal swelling and abdominal pain. An extremely severe hepatitis called fulminant hepatitis may occur in about 5% of patients with jaundice, which is characterised by fluid leaking into the abdomen, low blood protein content and clotting abnormalities, swelling of the brain and damage to erythrocytes causing haemolytic anaemia. There may also be difficulties with speech, swallowing and coordination, resulting in trembling, an unsteady walk and writing problems. A characteristic brown pigmentation in the cornea of the eye, called Kayser–Fleischer rings, may occur. Decreased renal functions and osteoporosis may occur prematurely in some patients. Associated psychiatric problems include severe insomnia, poor concentration, depression and suicidal impulses.

Wilson disease is diagnosed by a combination of blood and urine tests, eye examination and liver biopsy. A decrease in serum ceruloplasmin is seen in 95% of patients and an increased urinary excretion of copper is present in most, but not all, symptomatic patients. Kayser–Fleischer rings are visible with a slit-lamp examination performed by an optometrist or ophthalmologist in about 50% of patients presenting with liver disease. The definitive test is the demonstration of a high copper content in liver tissue obtained by biopsy.

Long-term maintenance therapy of Wilson disease is possible with D-penicillamine, trientine or zinc treatments. Penicillamine and trientine chelate copper, leading to an increased urinary excretion and reduced tissue levels. Zinc, with vitamin B_6 (pyridoxine) supplements, reduces copper absorption and promotes its loss from the GIT. Foods high in copper, such as shellfish (*Figure 10.24*), nuts, liver, chocolate and mushrooms, should be avoided, as should alcohol.

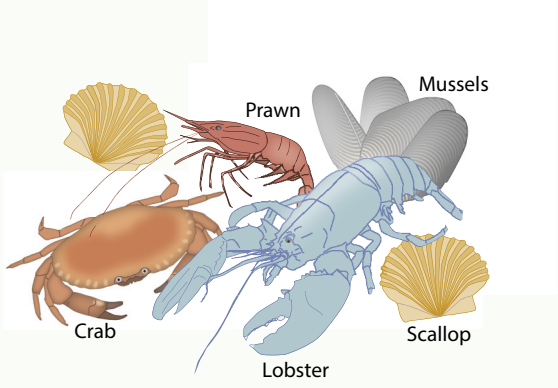


Figure 10.24 A selection of edible crustaceans and molluscs, colloquially called shellfish.

Medical therapies must be continued for life. Patients with acute liver failure or those with advanced liver disease who do not respond to medical therapy should be considered for liver transplantation, which effectively cures the condition and results in long-term survival of approximately 80% (*Chapter 6*). Genetic testing can be used to assist in counselling and to screen siblings of an identified patient.

Fluoride is necessary for the 'hardening' of bone and teeth. All the thyroid hormones contain iodine as described in *Chapter 7*.

Iron is a component of the prosthetic group haem, found in haemoglobin and myoglobin, where it maintains its oxidation state and binds a dioxygen molecule. Iron is also found in the haem of cytochromes (*Figure 10.25*) and in nonhaem iron proteins involved in electron transfer, where, of course, its oxidation state does alter.

Manganese is essential for the activities of a number of enzymes including pyruvate carboxylase and phosphoenolpyruvate carboxykinase that function in gluconeogenesis; arginase is a key enzyme of the urea cycle that detoxifies the ammonia produced by amino acid metabolism, and superoxide dismutase is a major antioxidant defence (*Figure 10.26*). Enzymes that contain molybdenum are common and catalyse several reactions in purine metabolism, for example xanthine oxidase, and those necessary for the synthesis of sex hormones. Selenium is a cofactor in, for example, glutathione peroxidases, and zinc in carbonic anhydrase and RNA polymerase.

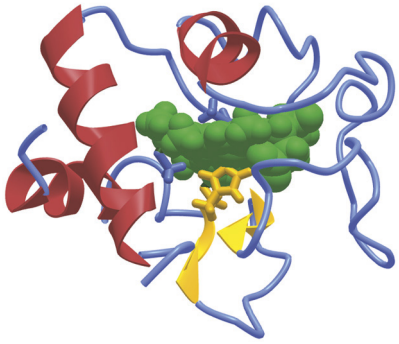


Figure 10.25 Molecular model of a cytochrome, with its haem group shown in green. PDB file 2N9I.

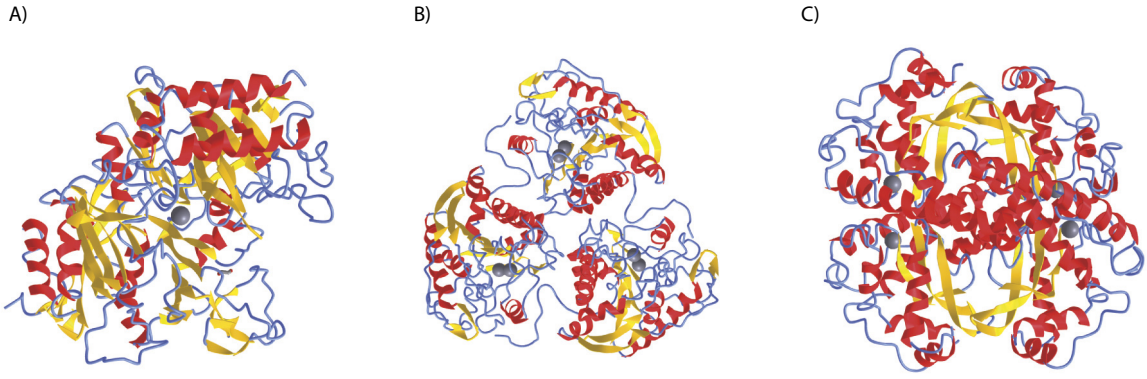


Figure 10.26 Molecular models of the manganese-containing enzymes (A) phosphoenolpyruvate carboxykinase, (B) arginase and (C) superoxide dismutase. In all cases, the dark grey spheres represent Mn ions. PDB files 1KHG, 2AEB and 1EM1, respectively.

10.3 NUTRITIONAL DISORDERS

Nutritional disorders can constitute major health issues affecting growth or causing specific diseases and even death. They can arise from an inadequate intake of food, such as protein–energy malnutrition; from an excessive food intake, for example obesity; or because of an inadequate amount of a specific dietary nutrient, such as scurvy in vitamin C deficiency, or toxic excesses of the vitamins A and D.

Protein–energy malnutrition

Protein–energy malnutrition (PEM) is the name for a spectrum of disorders that arise due to lack of food. Despite the name, affected individuals may be suffering not a lack of protein but a deficiency of dietary energy. Thus, dietary proteins that would normally be used for tissue repair or growth are used as a fuel. In addition, vitamin and mineral intakes are usually inadequate. In the developed world, PEM is rare and is usually associated with solitary elderly patients who are malnourished or with children suffering from neglect. Protein–energy malnutrition can also be life threatening in that it increases susceptibility to infectious diseases that would not normally be lethal. The clinical features of PEM vary depending upon the severity, from the merely underweight to two major, but distinguishable, conditions: **marasmus** and **kwashiorkor** (*Table 10.6*).

Condition	Expected weight for age (%)	Oedema
Underweight	60–80	Not present
Marasmus	<60	Not present
Kwashiorkor	60–80	Present
Marasmic kwashiorkor	<60	Present

Table 10.6 Distinguishing features of different types of protein–energy malnutrition.

Marasmus and kwashiorkor

Marasmus is a chronic disorder that develops over a period of months to years and is caused by an inadequate energy intake. It occurs in epidemics due to famine and is endemic in many areas of Africa, Asia and South America, and in patients with long-term illnesses, such as chronic pulmonary disease and anorexia nervosa (see later). Children with marasmus (*Figure 10.27*) fail to thrive, are emaciated and lack subcutaneous fat. **Cachexia**, that is, the muscle wastage associated with some chronic infections, such as tuberculosis (TB; *Chapter 4*), or the severe and prolonged weight loss seen in some cancers (*Chapter 17*), produces similar clinical features to marasmus but the aetiologies are different.

The name kwashiorkor is derived from one of the Kwa languages of Ghana and means 'the one who is displaced' and reflects the development of a nutritional condition in children typically three to five years old who have been abruptly weaned when a new sibling is born (*Figure 10.28*). Kwashiorkor is common in parts of the world affected by famine and where there is poor education and knowledge of nutrition. In the original description of kwashiorkor in children of The Gambia reported in 1932, Cicely Williams (1894–1992) implied that a deficiency of protein was its major cause, even when the energy input was adequate. Since then, a number of explanations for its development have been proposed. Environmental toxins, such as aflatoxins from mouldy foods, general conditions of overcrowding and poverty, a lack of other key nutrients and high rates of disease have all been implicated. Combined protein and energy deficiencies, although important, are not the key factors and it is generally accepted that the condition is likely to be due to deficiency of one of several nutrients, including copper, selenium, zinc and the vitamins folic acid, C, A, β carotene and E, which are associated with oxidative stress management. It is likely that infectious diseases are the precipitating factor because children with reduced antioxidant status exposed to the stress of an infection are most liable to develop kwashiorkor. Further, kwashiorkor usually occurs following infectious conditions, such as diarrhoea, or diseases, for example measles, indicating that its causes are not purely nutritional. Invading pathogens trigger a macrophage respiratory burst (*Chapters 4* and *5*) that considerably increases the total free radical load of the patient, and this may be the start of events that result in kwashiorkor. Unfortunately, its causes are still not fully known; siblings in the same household and on the same type of diet may develop marasmus or kwashiorkor. However, lower concentrations of the antioxidants β carotene, glutathione and vitamin E are observed in children suffering from kwashiorkor than in those affected by marasmus; both of course, have reduced levels compared with healthy children.

A significant feature of kwashiorkor is a large protruding abdomen due to oedema and an enlarged liver. The oedema is traditionally thought to occur because an inadequate intake of protein leads to a reduced plasma albumin concentration, which in turn causes oedema, although electrolyte disturbances, such as potassium deficiency and sodium retention, are also thought to play roles. The hepatomegaly occurs from a large infusion of fats into the liver, the cause of which is unknown. Early features of kwashiorkor include fatigue, irritability, lethargy, stunted growth, muscular wasting, oedema in the lower limbs and impaired neurological development. The skin may be affected by dermatitis and have areas of hypo- or hyperpigmentation, and the hair is thin, brittle, light-coloured and easily pluckable. The patients may also present with delayed wound healing and anaemia. If untreated, kwashiorkor can result in shock, coma and death, with a mortality as high as 60%. In the long term, it can lead to impaired physical and mental development.

Figure 10.27 Ethiopian children of the early 1980s showing signs typical of marasmus. Courtesy of the Catholic Fund for Overseas Development, London.

Margin Note 10.4
Kwashiorkor n the developed world

Kwashiorkor is rarely seen in developed countries, except in neglected children and the very elderly or as a consequence of nutritional ignorance. For example, cases of kwashiorkor have been reported in the USA in a small number of children fed somewhat idiosyncratic diets. Most of these cases were not associated with poverty and about half were associated with perceived or presumed food allergies. A significant portion of the food of some of these children were rice-based beverages, popularly referred to as 'rice milk', which contain less than half the amount of protein of breast milk. Other patients were fed brown rice emulsion, goat's milk or atole, a liquid emulsion of barley, water and sugar.

Figure 10.28 A child showing severe signs of kwashiorkor. Courtesy of the Catholic Fund for Overseas Development, London.

Obesity

Obesity is diagnosed when the patient's body mass index (BMI) is greater than 30 kg/m^2 (*Box 10.2*). It is characterised by an excess of fat in the body, particularly under the skin, and is generally recognisable when a person is 20% above their ideal body weight (IBW). Environmental and genetic factors can determine body weight, but overeating combined with a lack of exercise are its usual causes (*Box 10.3*). Very rarely, obesity may be secondary to endocrine disease (*Chapter 7*), for example in hypothyroidism, which decreases energy requirements; Cushing syndrome, where the distribution of body fat is altered; and some hypothalamic disorders that are associated with overeating. A number of diseases and clinical conditions occur more frequently in obese individuals (*Figure 10.29*), including type 2 diabetes mellitus (*Chapter 7*), CHD, hypertension (*Chapter 14*), cholelithiasis (*Chapter 11*) and osteoarthritis (*Chapter 18*). Not surprisingly, mortality is also greater as body weight increases. Some of the complications of obesity are listed in *Table 10.7*.

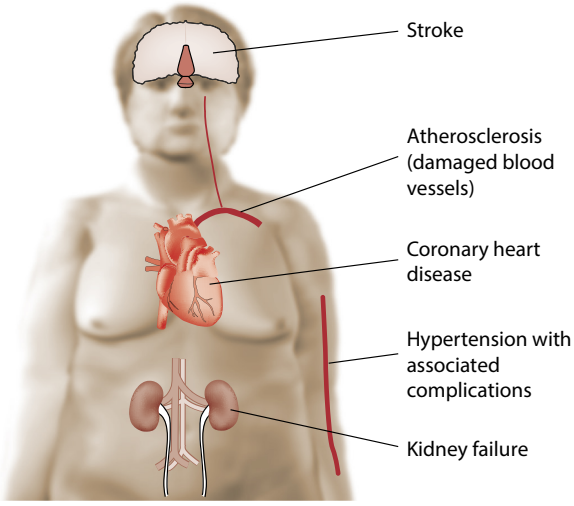


Figure 10.29 Schematic indicating the characteristic shape of an obese person with some of the associated complications indicated.

BOX 10.2 ASSESSING THE OVERWEIGHT AND OBESE

The generally accepted way to assess the weight of patients is to determine their **body mass index (BMI)**, where:

$$\text{BMI} = \text{weight}(\text{kg}) / \text{height}(\text{m})^2$$

Health risks associated with weight are lowest for a BMI of 20–25 kg/m^2. A BMI of less than 18.5 kg/m^2 is underweight and one of greater than 30 kg/m^2 is defined as obese. **Ideal body weight (IBW)** is another index that can be useful in assessing the nutritional status of a patient. This index is defined differently for men and women. The IBW of a man 5 feet tall is 106 pounds and this increases by an additional 6 pounds for each inch over the height of 5 feet. For a woman, the IBW at 5 feet tall is 100 pounds and this increases by 5 extra pounds for each additional inch.

Body mass index as a measure of obesity does not account for the distribution of excess weight in the body. It was suggested in the 1990s that central adiposity/obesity, that is, excessive abdomen fat, was a better predictor of the risks of developing clinical conditions, such as hypertension, non-alcoholic fatty liver, type 2 diabetes, obesity-related cardiovascular disease and strokes, than total obesity as measured by BMI (see *Table 10.7*). The most widely used methods of assessing central obesity are the waist-to-hip and, especially, the waist-to-height ratios (WHtRs). For both ratios, all measurements must be in the same units.

The waist/hip ratio is determined by measuring the circumference of the waist at its narrowest point and dividing it by that of the hip at its widest. Ratios greater than 0.80 for women and 0.95 for men are associated with an increased risk of developing the clinical problems associated with obesity. The WHtR is defined as the circumference of the waist divided by the height of the individual. In 2022, the National Institute for Health and Care Excellence (NICE) issued the following draft guidelines with respect to the WHtR:

- values of 0.4–0.49, that is, a waist measurement less than half the height, do not indicate any increased health risk
- values of 0.5–0.59 indicate an increased health risk and steps should be taken to reduce body weight
- values greater than 0.6 are associated with high central adiposity and are indicative of health risks and clinical intervention is indicated

These ratios are applicable to all ethnic groups with BMI values less than 35 kg/m^2, to adults and children over five years old and to persons of either sex. The NICE guidelines also warn that even people with healthy BMI values may have excessive WHtRs.

It is worth noting that measuring the circumference of the waist is inaccurate in those with a BMI greater than 35 kg/m^2, pregnant women and infants under two years old. Some have also suggested it is an inappropriate indicator for people who are very short or those older than 60 years who could have lost height with ageing (*Chapter 18*).

Disease/clinical disorder	Pathophysiology
Cholelithiasis	Increased cholesterol, bile stasis
Diabetes mellitus	Insulin resistance
Hypertension	Obesity, inappropriate lifestyle
Osteoarthritis	Increased wear and tear of joints
Respiratory disorder	Impaired lung ventilation
Vascular disease	Hypertension, dyslipidaemia, diabetes

Table 10.7 Clinical complications associated with obesity.

BOX 10.3 LEPTIN AND *OBESE/OBESE*

In 1994, the *obese** gene was discovered in mice. Its expression in adipocytes produces the hormone leptin, a polypeptide of 146 amino acid residues (*Figure 10.30*), which is secreted in amounts that are proportional to the mass of body fat. Among the effects of leptin are to decrease adiposity or total body fat and to pass through the blood–brain barrier and act on the hypothalamus to inhibit eating. It also stimulates the oxidation of fatty stores and prevents lipid accumulating in non-adipose tissues and so protects the individual against some of the adverse effects of fat accumulation, for example coronary artery disease.

Figure 10.30 Molecular model of a leptin molecule. PDB file 1AX8.

During periods when more energy is used than ingested, that is, the starved state, adipose tissue is metabolised and the amount of leptin secreted declines. Hence, the mouse is stimulated to eat more. When more energy is consumed than used, the converse holds. Mutations to the gene or its absence in mice can lead to excessive overeating and morbid obesity (*Figure 10.31*). Injections of synthetic leptin to obese mice cause weight loss. These findings led to great excitement

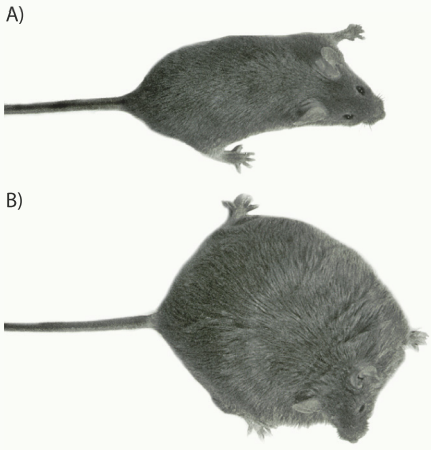


Figure 10.31 (A) Normal size and (B) genetically obese mice.

that obesity in humans could be a genetic phenomenon and controllable by leptin therapy. However, although the physiological effects of leptin are similar in humans and mice, mutations in the *OBESE** gene have been discovered in only a handful of obese humans. Two of these patients were members of the same highly consanguineous family and weighed 29 and 86 kg at the ages of two and eight years, respectively. Their *OBESE* genes had a deletion mutation of a single guanine nucleotide at codon 133 leading to a biologically inactive leptin. Injections of leptin relieved their symptoms. However, most obese people, in fact, secrete leptin in amounts that exceed its production in thinner people because they have a greater mass of fat. Unfortunately, they appear to be less sensitive to its effects.

* The convention is to write mice genes in lower case italic and human genes in upper case italic letters.

The increasing incidence of excessive body weight and obesity in adults, and particularly in children, in many developed countries is of increasing concern. For example, in 2021–2022, more than 26% of adults in England were obese and nearly 40% were overweight. Hence, only approximately 36% could be regarded as having a healthy weight. In 2006–2007, approximately 10% of children in England aged 10–11 years were obese. However, in 2022–2023, this had increased to over 22%, with approximately 14% being overweight and only about 62% having what could be regarded as a healthy weight (*Figure 10.32*). In 2019, it was estimated that obesity was costing the UK NHS £6.1 billion and wider society £27 billion.

Anorexia nervosa and bulimia nervosa

Anorexia nervosa (AN) and bulimia nervosa (BN) are both disorders of eating behaviour and body weight regulation. Anorexia is simply a lack of appetite; bulimia is derived, in part, from the Greek word *limos* meaning hunger. Both conditions have received considerable publicity in recent decades, although they were first reported many centuries ago. Women and men of any age can suffer from AN and BN, although the group at greatest risk is young females, with the conditions starting typically in the mid-teens. The global incidence of AN is estimated at 0.1–4% and 0–0.3% in females and males, respectively; comparable values for BN are 0.3–4.6% and 0.1–1.3%.

Anorexia nervosa is an extreme refusal or reluctance to eat, leading to severe weight loss. Associated psychological problems include compulsive exercising and the abuse

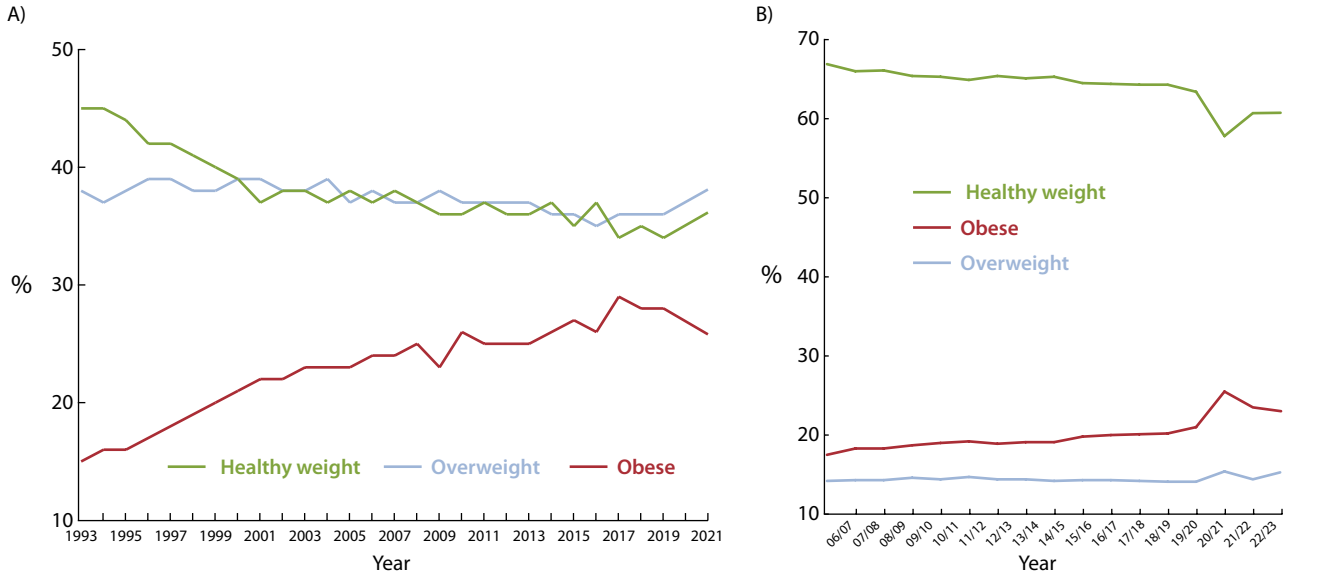


Figure 10.32 The increasing incidences of overweight and obesity in (A) adults and (B) children aged 10–11 years in England. The anomalous spike in obese and overweight children for 2020/2021 is attributed to the national lockdown and closure of schools due to the coronavirus disease 2019 (COVID-19) pandemic. Data courtesy of NHS Digital.

of laxatives and diuretics that often compound the reduced input of dietary energy. Patients are therefore normally extremely hungry and are obsessed with food but they still avoid eating, especially carbohydrates. The signs and symptoms of AN include a body weight at least 15% below the recommended weight, a 'wasted' appearance with reduced muscle mass and swelling of the joints. In younger females, puberty is delayed; an older female is likely to become **amenorrhoeic** and infertile because her body weight reduces to less than 45 kg and/or fat content becomes less than 22% of body weight. The skin is dry, hair is thin and nails are brittle. Decreased heart rate and blood pressure and constipation are common. If left untreated, long-term damage to the skeleton and cardiac systems are likely, and death can result from starvation, cardiac arrest or other complications. Sufferers of AN are often of above average intelligence but generally have a grossly dysmorphic view of their own bodies, seeing themselves as obese, even though they 'know' they are underweight. It is unclear as to what causes this view. It has been suggested that overreactions to relatively mild obesity, peer or social pressures regarding a perceived ideal human shape or a wish to delay the onset of menarche may all be linked to the condition.

Bulimia nervosa is characterised by episodes of excessive or 'binge' eating that induce feelings of guilt such that sufferers induce vomiting to void the food. This cycle of eating and induced vomiting can be repeated many times. As with anorexics, the use of laxatives, diuretics and dieting pills may be abused. The condition tends to affect older patients than those with anorexia. The signs and symptoms of BN include puffy cheeks due to enlarged salivary glands and often severely acid-damaged tooth enamel because of the excessive vomiting. Electrolyte imbalance, such as loss of potassium, can cause health problems and increases the risk of cardiac arrest (*Chapters 8* and *13*). However, patients generally manage to maintain their weight at an appropriate value and so the condition may not be noticed and can remain undetected for many years.

Disorders of vitamin nutrition

An inadequate dietary intake of a vitamin, its impaired absorption or insufficient utilisation of an adequate intake, and increased dietary requirements, for example in pregnancy, without a corresponding increased intake or an increased excretion of a vitamin give rise to **hypovitaminoses**. In many cases, the symptoms of a hypovitaminosis can be correlated with the known functions of the vitamin (*Section 10.2*), although in other cases they are rather generalised. Hypovitaminoses often develop over an extended period. Initially, there is depletion of body stores with a biochemical impairment, that is, a subclinical deficiency. This eventually results in an overt deficiency with frank signs and symptoms and is usually accompanied by other evidence of malnutrition, for

example PEM. A covert deficiency does not present with clinical features under normal conditions, but any trauma or stress may precipitate the hypovitaminosis. Starving individuals will suffer from multiple vitamin deficiencies. At the other extreme, an excess of the vitamin can be toxic and may result in a **hypervitaminosis**.

Clinical disorders associated with deficiencies of water-soluble vitamins

A deficiency of vitamin B_1 (thiamine) is relatively common in some developing countries because of the high consumption of foods such as polished rice, where the vitamin is lost during milling, and in chronic alcoholics who often have a poor diet. The consequences of a deficiency are depression, irritability, defective memory, peripheral neuropathy and beriberi. Beriberi, which literally means 'I cannot, I cannot' in Singhalese, occurs in two forms, which affect different body systems. Dry beriberi affects mainly the nervous system and wet beriberi affects the heart and circulation; both types usually occur in the same patient but one set of symptoms predominates.

Patients with the dry form may present late, with polyneuropathy and Wernicke–Korsakoff syndrome (*Box 10.4*). The polyneuropathy is characterised initially by heavy, stiff legs, then weakness, numbness and paraesthesia and absent ankle jerk reflexes. Later stages involve the trunk and arms. Wet beriberi, also known as shoshin from the Japanese for acute heart damage, is less common and characterised by oedema. Symptoms appear rapidly with acute heart failure in addition to the polyneuropathy. It is highly fatal and known to cause sudden deaths in young migrant labourers in Asia whose diet consists largely of white rice.

Vitamin B_2 (riboflavin) deficiency is rare in the developed world and usually only seen in alcoholics who normally have diets lacking other nutrients or suffer malabsorption.

BOX 10.4 WERNICKE KORSAKOFF SYNDROME OR CEREBRAL BERIBERI

Wernicke–Korsakoff syndrome or cerebral beriberi, also known as Wernicke disease or Korsakoff psychosis, is a brain disorder with the loss of specific functions. The term *Gayet disease* is applied when lesions are more extensive than those in the Wernicke type. Incidence and prevalence data for Wernicke–Korsakoff syndrome are poorly reported, although one study has indicated a prevalence of 1 to 3%. The condition arises primarily from a deficiency of vitamin B_1 (thiamine) or secondary to alcoholism and/or starvation. Even in alcoholics (*Chapter 12*) on a balanced diet, which is very unusual, the heavy drinking interferes with the absorption of thiamine from the gastrointestinal tract (GIT).

The syndrome includes two separate conditions. The first, Wernicke encephalopathy, involves damage to nerves of the central and peripheral nervous systems. The second, Korsakoff syndrome or Korsakoff psychosis, begins as Wernicke symptoms diminish and involves damage to areas of the brain involved with memory. Patients often attempt to hide the memory loss by confabulating, that is, telling detailed and credible stories that are untrue, although these are not usually deliberate attempts at deception because the patient often believes them to be factual. Other symptoms include double vision and other eye abnormalities, a loss of muscle coordination, unsteady gait and hallucinations. Many of these symptoms are indicative of alcohol withdrawal and may also be present or develop even when Wernicke–Korsakoff syndrome is not present. Other disorders related to the abuse of alcohol may also be apparent.

Abstinence from alcohol or moderate intake together with adequate nutrition reduces the risk of developing Wernicke–Korsakoff syndrome. Thiamine supplements and a good diet may also help prevent the condition occurring in a heavy drinker but not if damage has already occurred.

The syndrome is diagnosed following an examination of the nervous and muscular systems that demonstrates polyneuropathy or damage to multiple nerves and decreased or abnormal reflexes. Testing of gait and coordination can indicate damage to the parts of the brain that control muscle coordination. Muscles may be weak and atrophied and abnormalities of eye movement may be present. The patient may present as **cachectic**. Clinical tests should demonstrate a low serum vitamin B_1 concentration and low erythrocyte transketolase activity but an increase in pyruvate concentration. A magnetic resonance image or computed tomography scan of the brain (*Chapter 18*) of a Wernicke–Korsakoff syndrome patient can show changes in the thalamus or hypothalamus. If the case history is significant for chronic alcohol abuse, then serum or urine alcohol concentrations and liver enzyme activities may be increased above normal.

Wernicke–Korsakoff syndrome is a life-threatening condition with a mortality rate of 10 to 20%. Treatment is aimed at controlling the symptoms as much as possible and preventing any progression of the disorder, although some symptoms, particularly the loss of memory and cognitive skills, may be permanent. Hospitalisation is required for the initial control of symptoms. In lethargic, unconscious or comatose patients, appropriate monitoring and care is necessary, particularly to prevent obstruction of the airways. Carbohydrate loading, for example by glucose infusion, may precipitate Wernicke encephalopathy in at-risk patients but can be prevented by supplements of vitamin B_1 prior to the glucose infusion. Injections or oral supplements of vitamin B_1 may improve some symptoms but do not generally lead to the recovery of memory and intellect. Total abstinence from alcohol is required to prevent progressive damage to the brain and peripheral nerves, and a well-balanced, nourishing diet is recommended. Joining a support group whose members share common experiences and problems can often help in coping with the stresses of the illness. There may be a need for custodial care if the loss of cognitive skills is severe.

However, it especially affects children in developing countries in Asia and Africa. Other susceptible people include the elderly with a poor diet, regular laxative users and vegan athletes, who are at risk of B_2 deficiency. The main metabolic effects are on lipid metabolism, which leads to clinical features of weakness, dermatitis, glossitis, insomnia and sensitivity to bright light, and normochromic and normocytic anaemia (*Chapter 14*). A deficiency of vitamin B_2 is rarely fatal because it is present in most foods, which provide adequate amounts to maintain function, and there is efficient reutilisation of the vitamin during the turnover of flavoproteins, so that little is lost. A deficiency of vitamin B_3 (niacin, nicotinic acid) causes pellagra, where the patients present with weight loss, anaemia, dementia, dermatitis and diarrhoea. Pellagra can be primary when the diet is deficient in B_3, such as when maize is the staple, but it can be secondary, where other diseases interfere with absorption. Causes of secondary pellagra include prolonged diarrhoea, cirrhosis of the liver, alcoholism (*Chapter 12*) and use of isoniazid (*Figure 10.33*, but see also *Figure 3.47*) to treat TB.

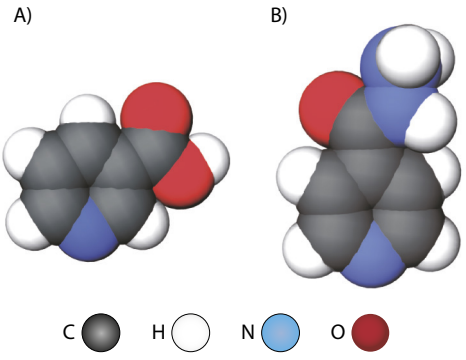


Figure 10.33 Computer drawn models of (A) niacin and (B) isoniazid. ⏎

A deficiency of vitamin B_5 (pantothenic acid) is rare but has been induced experimentally in animals by feeding a diet virtually devoid of the vitamin or by giving metabolic antagonists, such as ω methylpantothenic acid. Symptoms included irritability, fatigue, malaise, GIT problems, muscle cramps and paraesthesia. Historically, pantothenic acid deficiency has been implicated in the 'burning feet' syndrome experienced by severely malnourished prisoners of war.

A deficiency of vitamin B_6 impairs the synthesis of γ aminobutyrate and amino acid metabolism and may be implicated in the development of some cancers (*Chapter 17*). Clinical features associated with a deficiency of vitamin B_6 include peripheral neuropathy. However, deficiencies are uncommon because it is widely available. Cases have been reported in infants fed overheated and inadequately fortified milk formula who developed severe symptoms: irritability, **opisthotonos** (which is an arching of the head, neck and spine backwards) and convulsions that are relieved by vitamin B_6 supplements.

Moderate deficiencies of vitamin B_6 can occur in patients treated over long periods with isoniazid (*Figure 10.33*) and penicillamine (*Chapter 3*) that can form biologically inactive complexes with pyridoxal. This can lead to abnormalities in methionine and tryptophan metabolism. Severe deficiency is vanishingly rare but can lead to convulsions in babies, which are readily resolved with B_6 administration.

Patients suffering a lack of only vitamin B_7 (vitamin H, biotin) are extremely rare. They tend to be individuals who consume extremely large amounts of raw eggs, since the white contains the protein avidin, which binds biotin tightly (*Figure 10.34*). The vitamin is not released unless the egg is cooked; however, more than 20 raw eggs must be consumed daily for this to become a serious problem. Biotin deficiency occasionally occurs in hospital patients on parenteral nutrition and leads to anorexia, nausea, dermatitis and depression.

Vitamin B_9 (folic acid) deficiency is relatively common and occurs because of an inadequate dietary intake in alcoholics and the elderly or because of increased requirements, for example during pregnancy and in diseases associated with increased cell and nucleic acid turnover, such as leukaemia. The deficiency can also arise because some drugs affect folate metabolism, for example anticonvulsant drugs such as phenobarbitone, phenytoin and primidone, or, like phenytoin and carbamazepine, result in its malabsorption. The major clinical feature of folic acid deficiency is megaloblastic anaemia (*Chapter 13*). An increased intake of folic acid before and during pregnancy is associated

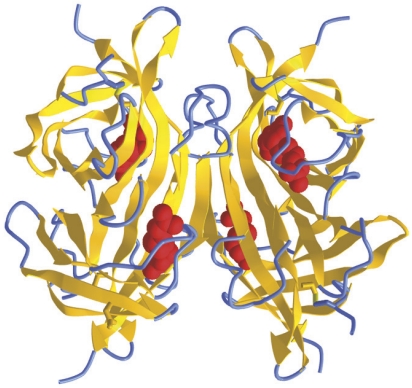


Figure 10.34 Molecular model of vitamin B_7 (biotin, vitamin H), shown in red, bound to avidin. PDB file 1AVD. ⏎

with a decreased risk of the fetus developing neural tube defects. Taking supplements of folic acid before conception and in the first three months of pregnancy reduces the incidence of spina bifida by more than 20%. In women who have already had a pregnancy affected by a neural tube defect, taking a daily folic acid supplement reduces their chances of a similarly affected pregnancy by approximately 70%. Folic acid supplements are associated with decreased amounts of homocysteine in the plasma, with a reduced risk of cardiovascular disease (*Chapter 14*).

A deficiency of vitamin B_{12} (cobalamin) causes megaloblastic anaemia (*Chapter 13*) and degeneration of the spinal cord. However, deficiencies are generally uncommon, even in cases of severe malabsorption (*Chapter 11*), since considerable amounts of the vitamin are stored in the liver. Deficiencies can occur in vegans, since dietary sources are animal products. When a deficiency does occur, it is seen most commonly in pernicious anaemia, an autoimmune disease characterised by the lack of intrinsic factor required for absorption of the vitamin from the GIT (*Chapters 11* and *13*). Bacterial overgrowth in the GIT or diseases affecting the small intestine, such as Crohn disease (*Chapter 11*), can also lead to a deficiency. Elderly people with atrophic gastritis absorb vitamin B_{12} poorly. However, the first sign of this is neuropathy because the anaemia is hidden by folic acid intakes.

A lack of vitamin C (ascorbic acid) causes scurvy. Inadequate dietary intake occurs in infants aged 6 to 12 months who receive processed milk without fruit juices or vitamin C supplements and in the elderly who have vitamin C deficient diets. The clinical features of scurvy include skin **papules** and **petechial haemorrhages** (*Chapter 6*), muscle haemorrhages, poor wound healing, gum disease, anaemia and osteoporosis. It has been suggested that large (gram) doses of vitamin C reduce the incidence, duration and severity of the common cold without any side effects, but there is no scientific evidence for these claims.

Clinical disorders associated with deficiencies of fat-soluble vitamins

A deficiency of vitamin A is rare in the developed world; indeed, large amounts are stored in the liver and transported around the body bound to retinol-binding protein in the plasma (*Figure 10.35*). In contrast, deficiencies are major health issues in parts of Africa and South-East Asia. In pregnant women, vitamin A deficiency not only causes night blindness but may also increase the risk of maternal mortality. Affected patients may present with anaemia and disorders of ectodermal tissues and increased risk of disease and death from severe infections.

A deficiency of vitamin D impairs mineralisation of bone matrix causing rickets in children and osteomalacia in adults. Overt rickets is rare in the UK, although subclinical rickets affects about 10% of young children. Rickets results in deformities of the legs, which bow owing to the weight of the body (*Figure 10.36*), skull, ribcage and pelvis. During the 1950s, rickets was virtually eradicated in the UK by fortification of infant foods with vitamin D. Unfortunately, some susceptible children suffered vitamin D toxicity and developed hypercalcaemia (*Chapter 8*). As a consequence, the amount of vitamin D added to foods was reduced and rickets reappeared. Osteomalacia causes pain in the bones, especially of the pelvis and legs, and there is an increased susceptibility to fractures of the long bones following minor trauma. It is not uncommon in the UK, especially among women of some ethnic origins because of their lack of exposure to sunlight during purdah and the wearing of traditional clothes.

Vitamin E deficiency is rare. When it does occur, it is most likely in newborn babies because transfer of vitamin E across the placenta is poor and babies also have less adipose tissue, where most of the vitamin is stored. Deficiencies of vitamin E can occur with long-term parenteral nutrition and in prolonged and severe steatorrhoea (*Chapter 11*). Vitamin E deficiency in children causes irritability, oedema, haemolytic anaemia and neurological dysfunction. Decreased concentrations of plasma vitamin E have been associated with progression of atherosclerosis and growth of some tumours. Other symptoms such as ataxia, dysarthria, sensory loss and paraesthesia have been described in adults.

The clinical effects of vitamin K deficiency include a prolonged clotting time and a bleeding tendency as described in *Chapter 13*. A maternal deficiency can lead to serious bone defects in the fetus. Vitamin K deficiency is most common in neonates because the vitamin is unable to cross the placenta and milk is a poor dietary source, so the baby is at risk of bleeding. Newly born babies lack bacteria in their GIT that can synthesise vitamin K, although the significance of intestinal biosynthesis by bacteria is debatable anyway. For this reason, it has been recommended that all neonates are given a single prophylactic dose of vitamin K. In adults, a deficiency may be seen in fat malabsorption or in people using antibiotics that reduce intestinal bacteria (see *Margin Note 10.5*). An

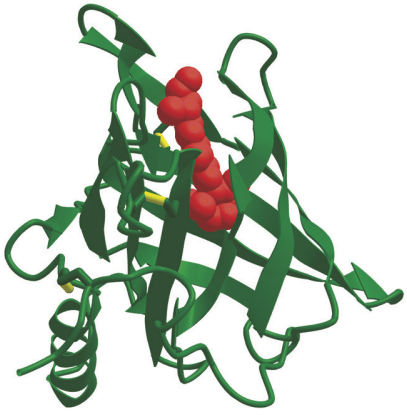


Figure 10.35 Molecular model of serum retinol-binding protein with a bound retinol form of vitamin A (red). PDB file 1BRP. ⏎

Figure 10.36 A child with the characteristic bowed legs of rickets. ⏎

inadequate intake may reduce the density of bone and increase the risk of osteoporosis and associated fractures, particularly in postmenopausal women.

Clinical disorders associated with hypervitaminoses

Hypervitaminoses are relatively rare compared with deficiencies. An excess of vitamins B_3 and B_6 can be toxic and is usually associated with excessive intakes of vitamin supplements. Large doses of niacin are associated with a variety of clinical problems, including abnormalities of liver functions, hyperglycaemia, vasodilation and an increase in plasma uric acid concentration. Daily doses of vitamin B_6 greater than 500 mg over an extended period can cause a sensory neuropathy.

Most examples of hypervitaminoses are associated with vitamins A and D. Vitamin A is stored in the liver, and excessive dietary intake over prolonged periods can lead to a toxic overload. Typical symptoms are pain in the bones, a scaly dermatitis, nausea and diarrhoea with enlargements of the liver and spleen. Most cases of vitamin A toxicity are caused by patients overdosing with vitamin supplements. The only natural food known to contain dangerous levels of the vitamin is polar bear liver; not a common dietary item in most supermarkets or, indeed, societies (*Figure 10.38*)! Animal studies have shown that vitamin A can produce teratogenic effects when administered in high doses. The consumption of excessively large amounts of vitamin A during pregnancy may increase the risk of congenital malformations.

Excess of vitamin D is, again, largely associated with the overconsumption of vitamin supplements. Toxicity is due to overstimulation of calcium absorption from the gut and excessive resorption from bone, which results in its demineralisation. The weakening of the bone and hypercalcaemia (*Chapter 8*) promote metastatic calcification and a tendency in the patient to form kidney stones.

Nutritional disorders of minerals and trace elements

Clinical disorders arising from deficiencies in the dietary intakes of minerals are not uncommon and a number are described in *Chapter 8*, where conditions caused by excessive mineral ingestion, which are less common, are also outlined. The total quantity of any one trace element in the body is usually less than 5 g and these elements are often required in quantities of less than 20 mg per day; hence, dietary deficiencies are uncommon.

Chromium deficiency can occur in patients on parenteral nutrition without adequate supplementation and leads to glucose intolerance. An excess of chromium (II) has no known symptoms, although chromium (III) and especially chromium (VI) compounds are toxic. A deficiency of cobalt is rare and causes indigestion, diarrhoea, weight loss and a loss of memory. An excess of cobalt is not associated with any known symptoms, although a high intake over a prolonged period may lead to infertility in men. In addition, there are occasional reports of cobalt cardiomyopathy following occupational exposure.

Copper deficiency is uncommon, except in patients on synthetic oral or on long-term parenteral nutrition. It can occur in infants because of malnutrition, malabsorption, chronic diarrhoea or prolonged feeding with low-copper milk diets. Premature infants are particularly susceptible because of their low copper stores in the liver. Copper deficiency causes neutropenia and hypochromic anaemia in the early stages, both of which respond to dietary copper but not iron. This is followed by bone abnormalities such as osteoporosis, decreased pigmentation of the skin, pallor and neurological abnormalities in the later stages. A dietary excess of copper is rare but occasionally happens following food contamination and causes salivation, stomach pain, nausea, vomiting and diarrhoea (see *Box 10.1*).

Nearly half of dietary fluoride is taken up by bone and can influence its mechanical properties. The incorporation of fluoride in tooth enamel as fluorapatite makes teeth more resistant to dental caries, as it resists breakdown by acids. Fluoride deficiency leads to impaired bone formation and dental defects (*Box 10.5*).

Iodine deficiency causes an enlargement of the thyroid gland called a goitre (*Figure 10.39*), while a prolonged deficiency in children can lead to cretinism (*Chapter 7*). Iodine deficiency is most common in upland areas with thin limestone soil because minerals, such as iodine, are easily leached from these types of soils, and plants in the area are iodine deficient. The situation is exacerbated if commercial iodised salt is not available. In some areas of Africa, Brazil and the Himalayas, more than 90% of the population may develop goitre due to iodine deficiency. Excessive intake of dietary iodide can also lead to goitre. This is most commonly seen in Japanese communities with high dietary intakes of seaweed.

Margin Note 10.5 ⏎
Warfarin and vitamin K

Warfarin is often used therapeutically as an anticoagulant (*Chapter 13*). However, because of the similarities in structures, it is a vitamin K antagonist (*Figure 10.37*). Patients on warfarin therapy require increased amounts of the vitamin.

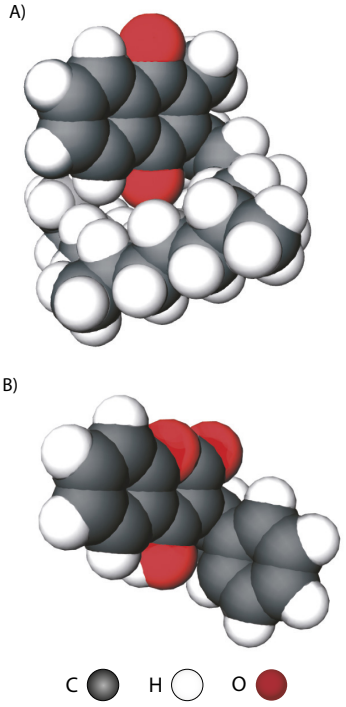


Figure 10.37 Computer drawn models of (A) vitamin K_1 and (B) warfarin. ⏎

Figure 10.38 A polar bear (*Ursus maritimus*). ⏎

BOX 10.5 FLUORIDE AND DENTAL CARIES

The bulk of teeth are composed of the hard tissue dentine (see *Box 11.2*). The exposed part of the tooth or crown is covered with enamel. Dental caries or tooth decay is caused by bacteria in dental plaque and is the most common cause of toothache (*Chapter 3*). It is particularly common in children. Plaque is a thin layer of mucins and other complex molecules that often contains bacteria. These can catabolise sugars, producing organic acids that can demineralise the enamel and the underlying dentine. The development of dental caries depends upon a balance between the plaque and bacteria and the ability of enamel to resist attack. The most common cause of dental caries is poor dental hygiene and a frequent intake of dietary sugars. Regular brushing of the teeth can remove plaque and reduce acid production since it takes about 24 hours for the bacteria to reestablish themselves. Treatment of caries involves removing the decayed part of the teeth by drilling and replacing the decayed material with a filling.

Epidemiological studies have demonstrated the benefits of adding fluoride to drinking water (fluoridation) because as little as one part per million (ppm) can prevent dental caries. Studies in the UK showed that children living in areas with water fluoride concentrations of about 1 ppm had 50% fewer cases of dental caries than those in areas where the water fluoride was below 0.1 ppm. Indeed, removal of fluoride from the water increased the incidence of caries. These studies have been supported by others worldwide, where similar results have been obtained despite ethnic, social, climatic and dietary differences. Fluoride in toothpastes, in mouthwashes or when consumed in a liquid or tablet form is also effective in preventing dental caries.

The addition of fluoride to drinking water at concentrations greater than 12 ppm promotes dental **fluorosis**, a mottling of the teeth as they form in the jaws. Its effects are generally cosmetic and only cause functional problems when the disorder is severe. Another concern of fluoridation is its possible association with arthritis. There is, however, a weak association between fluoridation and an increased susceptibility to bone fractures.

The addition of fluoride to drinking water has also raised ethical issues because it is seen by some as an infringement of personal liberty in that individuals have no choice but to drink water containing fluoride. The use of fluoridation requires careful consideration of the ethical issues and the balance between its beneficial and any potentially harmful effects.

A lack of iron leads to anaemia, which is described in *Chapter 13*. Increased intake of dietary, medicinal or transfused iron can cause haemosiderosis, characterised by deposits of iron compounds in organs such as the liver and heart (see *Figure 13.5*). Acute ingestion of large doses of ferrous salts can be fatal, particularly in children under the age of two.

Manganese deficiency is rare but is believed to cause ataxia, hearing loss and dizziness. Inhalation of dust from mining and other industrial sources can lead to manganese toxicity and causes severe neurological dysfunction, similar to Parkinson disease (*Chapter 18*). Concentrations of manganese in food and water must exceed 100 mg/kg body weight to produce manganese toxicity. Such high concentrations are thought to stimulate xanthine oxidase activity, leading to increased plasma uric acid concentrations and gout (*Chapter 8*). However, few data are available, but the major signs include diarrhoea and anaemia.

A dietary deficiency of molybdenum has not been reported. However, molybdenum deficiency was noted in a patient on parenteral nutrition who suffered from mental disturbances that progressed to coma. Supplementation of molybdenum improved the patient's clinical condition.

Selenium deficiency results from a low dietary intake in parts of the world with soils of low selenium content and has been reported in patients on long-term parenteral nutrition. A deficiency of selenium may lead to a cardiomyopathy. In China, this is called Keshan disease and affects young women and children in selenium-deficient regions. Selenium deficiency is a basic factor in Keshan disease; however, its occurrence is seasonal and it is associated with viral infections that cause inflammation of the heart. Prophylactic selenium prevents the disease developing, but selenium supplements do not reverse heart muscle damage. A large intake of selenium causes selenosis, a condition characterised by loss of hair, skin and nails.

Zinc deficiency is relatively common in populations in rural areas of the Middle East and subtropical and tropical areas where unleavened whole wheat bread can provide up to 75% of the energy intake. The little zinc in wheat is bound by the relatively large amounts of the metal ion chelator phytate (*Figure 10.40*) and fibre present, which inhibit its absorption. This is not a problem in leavened bread, as yeasts produce phytases that inactivate the phytase. Zinc deficiency also occurs during prolonged parenteral nutrition if inadequate amounts are provided. A deficiency is associated with a period of severe catabolism, as would occur in PEM. A severe deficiency occurs in the skin condition acrodermatitis enteropathica, where there is an inherited defect in GIT zinc absorption. A deficiency of zinc delays the onset of puberty. High doses of zinc reduce the amount of copper the body can absorb, causing anaemia and weakness of the bones.

Figure 10.39 A well-developed goitre. Courtesy of S.J. Nixon, Royal Infirmary of Edinburgh, Scotland.

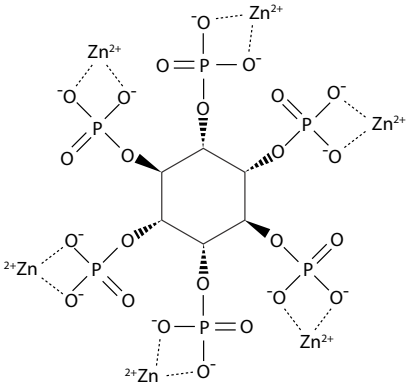


Figure 10.40 The chelation of Zn^{2+} by phytate.

10.4 INVESTIGATING NUTRITIONAL DISORDERS

The main objectives in investigating nutritional disorders are to:

- detect malnutrition
- determine the most appropriate treatment
- monitor patient progress during treatment

A number of fairly general investigations are available for patients suspected of suffering from malnourishment, for example taking a medical history, physical examination and laboratory investigations. A patient's medical history may help to reveal the underlying basis of the complaint. A significant part is determining the dietary history, which usually involves dietary recall of the last one to seven days, or food frequency questionnaires where patients estimate how often a particular food on a list has been eaten. A medical history also includes specific questions concerning recent weight changes, dietary habits (for example alcohol consumption), appetite, GIT function, use of vitamin and mineral supplements, and drug(s) intake. A physical examination can include anthropometric measurements, such as weight, height, BMI (*Box 10.2*), arm circumference and skinfold thickness. Arm circumference is an indicator of muscle mass. The arm circumference is measured at the midarm point in millimetres (*Figure 10.41*): values are typically 293 mm in males and 285 mm in females. Skinfold thickness indicates the extent of subcutaneous fat stores. It is commonly measured at the biceps, triceps, and subscapular and suprailiac sites in both males and females. In, for example, the triceps measurement, the arm is allowed to hang freely and a fold of skin above the midpoint of the arm is pulled from the underlying muscle and measured to the nearest millimetre using callipers (*Figure 10.41*). Values are typically 12.5 and 16.5 mm for males and females, respectively. Measurements at the four sites should be made in duplicates that should not vary by more than 1 mm. If consecutive measurements become increasingly smaller, the fat is being compressed and other sites should be assessed while the tissue recovers. However, the accuracy of the test depends upon selecting the sites correctly, an appropriate technique in taking the measurements and the experience of the tester; hence, the trend is more important than the values at any one time. Values for arm circumference and skinfold thickness are compared with standard values for the same sex. Undernutrition is indicated if they are less than 90% of these values. The appearance of the skin, hair, nails, eyes, bones, teeth and mucous membranes may reveal signs of malnutrition. A number of general laboratory-based investigations on blood and urine samples can aid in diagnosing malnutrition (*Table 10.8*). These tests are not reliable when used in isolation, as their findings are also altered in some disease states. Hair and nail samples are of questionable clinical value and are not routinely taken.

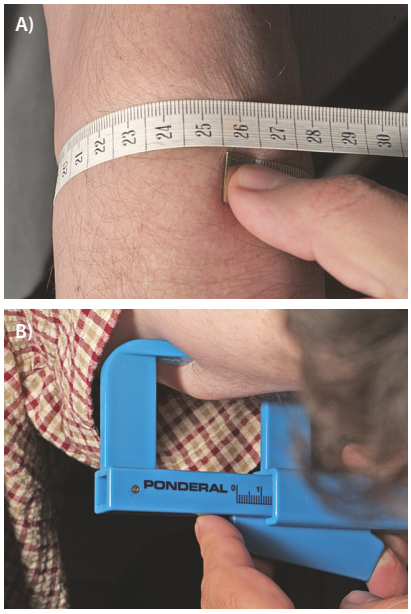


Figure 10.41 Measuring (A) arm circumference and (B) skinfold thickness. ⏎

Test	Clinical value in assessment of malnutrition
Total lymphocyte count (*Chapter 13*)	Reduced in protein–energy malnutrition, especially kwashiorkor
Serum albumin concentration	Reduced in kwashiorkor, poor protein intake and zinc deficiency
Serum creatinine concentration (*Chapters 1* and *8*)	Reduced in muscle wasting because of energy deficiency
Haemoglobin (*Chapter 13*)	Reduced in anaemia possibly due to nutritional deficiency
Total iron-binding capacity (*Chapter 13*)	Reduced in kwashiorkor and poor protein intake
Prothrombin time (*Chapter 13*)	Increased in vitamin K deficiency

Table 10.8 Laboratory tests used in investigating malnutrition. ⏎

Protein–energy malnutrition

A diagnosis of PEM must identify the type present. Individuals affected by marasmus present with decreased anthropometric measurements, decreased body weight and loss of muscle mass. There is also a decline in body temperature, pulse and metabolic rate. Serum protein levels are often low but can be normal. Children suffering from PEM have reduced rates of growth and are shorter, particularly in marasmus, compared with their normal counterparts. This can be easily assessed using charts of growth rates

(*Figure 10.42*). They also have an impaired immune system because of reduced protein synthesis, particularly of immunoglobulins; as a consequence, infections such as measles that a child would normally be expected to survive are common causes of death in severe cases. A diagnosis of kwashiorkor is made following a thorough physical examination together with a medical and dietary history. Patients affected with kwashiorkor have the characteristic swollen abdomen, show hypoalbuminaemia and a reduced lymphocyte count or **lymphopenia**.

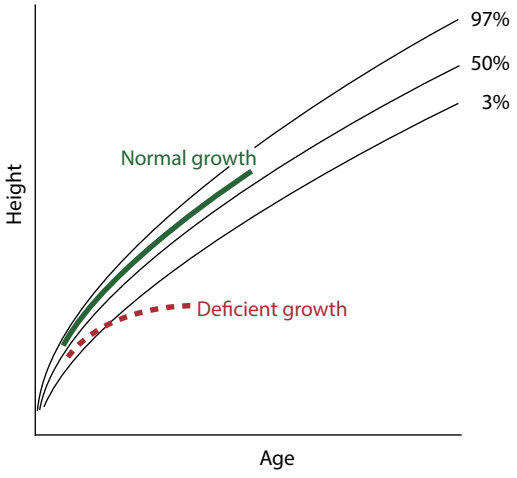


Figure 10.42 A generic growth chart. Charts are available for different sexes and ethnic groups.

Obesity

In addition to arm circumference and skinfold thickness (see above), the distribution of fat in obese patients can be assessed by determining the waist-to-hip ratio and/or the WHtR (see *Box 10.2*). Ultrasonography can also be used on soft-tissue adipose material. The technique reflects ultrasound waves from internal tissues or organs and the echo patterns are analysed to form a picture called a **sonogram** of body tissues. The extent of fat deposition can be accurately determined but is more expensive than other methods and not widely used.

Anorexia nervosa and bulimia nervosa

Anorexia nervosa and BN are difficult to identify, given that the patients often take extreme measures to hide the conditions and, indeed, often refuse to admit to any problem. Diagnosis relies on a medical history and clinical investigation supported by anthropometric measurements, coupled with a sympathetic psychological assessment.

Vitamins

The majority of vitamin disorders encountered in clinical practice are deficiencies. The investigative procedures are varied and depend upon the vitamin in question. Chemical tests can help to confirm the diagnosis of overt vitamin deficiencies and may enable diagnoses to be made at a relatively early stage. The types of tests used include direct measurements of the concentration of the vitamin or one of its metabolites in plasma, serum, erythrocytes, urine or tissue biopsies. The concentration of vitamin in plasma does not necessarily reflect body vitamin status and a measurement of the concentration in blood cells may be a better indicator. Enzyme-based tests are available for some vitamins. Metabolites that accumulate in the blood or urine following the blockage of a metabolic pathway normally catalysed by an enzyme that requires a vitamin to form a cofactor or coenzyme may also be investigated.

Water-soluble vitamins

Vitamin B_1 (thiamine) deficiency can be assessed by direct measurement of its concentration in plasma or indirectly by determining the increase in erythrocyte transketolase activity in the presence of added TPP (*Figure 10.43*). The increase in activity is called the activation coefficient. A coefficient less than 15% is considered normal; an increase of 15–25% indicates a marginal deficiency, while increases greater than 25% with clinical signs are indicative of severe thiamine deficiency. Thiamine deficiency can also be assessed by the clinical response to administered thiamine, that is, an improvement in

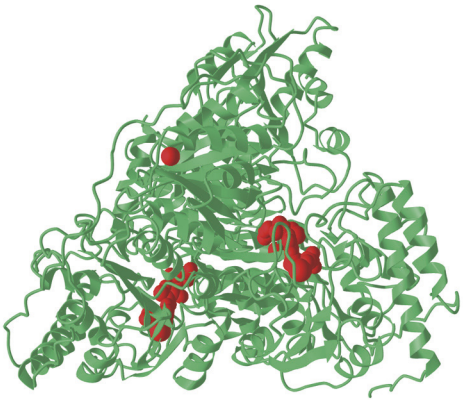


Figure 10.43 Molecular model of a transketolase with its bound thiamine pyrophosphate (TPP) and Ca^{2+} shown in red. PDB file 3MOS.

the condition after administering thiamine supplements. The nutritional status of vitamin B_2 (riboflavin) is investigated in a similar manner by determining the activation of glutathione reductase activity of erythrocytes in the presence of added FAD (*Figure 10.44*).

Assessing the nutritional status of vitamin B_3 (niacin) is more difficult. The usual method is to determine the concentrations of metabolites of niacin, for example 1-methylnicotinamide and 1-methyl-3-carboxamido-6-pyridone, in urine samples. Both are reasonably good measures of niacin status, as is the ratio of the concentrations of NAD^+ to $NADP^+$ in erythrocytes. A ratio of less than 1.0 may identify subjects at risk of developing a niacin deficiency. The nutritional status of vitamin B_5 (pantothenate) can also be assessed by determining its concentration in plasma or urine samples. In general, plasma pantothenate concentrations decrease in patients on a pantothenic acid deficient diet. However, its concentration in blood responds less readily to intake than the concentration in urine.

The status of vitamin B_6 (pyridoxal, etc.) can be investigated in manners similar to those for thiamine and riboflavin by determining the activation coefficients of erythrocyte alanine and aspartate transaminase activities (ALT and AST, respectively) in the presence of the cofactor pyridoxal phosphate. Alternatively, vitamin B_6 status may be assessed by the tryptophan loading test (*Figure 10.45*). However, the activity of kynureninase decreases markedly in B_6-deficient patients. If such a patient is given an oral dose of 50 mg/kg body weight of tryptophan, then there is an increase in the amounts of kynurenate and xanthurenate formed and these appear in the urine. Generally, less than 30 mg of xanthurenate is excreted daily; higher amounts are indicative of vitamin B_6 deficiency. However, some other disorders of tryptophan catabolism can also lead to an increase in xanthurenate production, so abnormal results must be treated with caution.

A possible deficiency of vitamin B_7 (vitamin H, biotin) can be investigated by measuring its concentration in whole blood, serum or urine. Determining plasma biotin is not a reliable indicator of status. Changes in urinary excretion of biotin or of its metabolites are better indicators of biotin status.

Folate status may be investigated by directly measuring its concentration in serum samples or erythrocytes, although these tests are associated with a number of problems. Serum folate tends to reflect dietary intake over the previous few weeks. Patients with both acquired and inherited folate deficiencies may remain moderately deficient for months or years, taking in just enough folic acid to prevent low erythrocyte folate concentrations and frank anaemia. Erythrocyte values are not sensitive to short-term variations; depletion occurs only in the later stages of deficiency and is usually accompanied by megaloblastic anaemia. Both erythrocyte and serum folate studies must be performed. Severe folate deficiency is accompanied by a macrocytic anaemia (*Chapter 13*), although the size of erythrocytes may be entirely normal in lesser degrees of depletion. Serum vitamin B_{12} (cobalamin) concentrations should also be measured when evaluating folate deficiency, since if either vitamin is deficient it can lead to a failure in absorption by megaloblastic intestinal cells resulting in a secondary deficiency of the other. Formiminoglutamic acid (FIGLU) is a substrate for the folate-dependent enzyme formimino-glutamate formiminotransferase, which is required for histidine catabolism. A deficiency of folate allows FIGLU to accumulate and it is excreted in urine, providing a sensitive test of deficiency. However, FIGLU also increases in vitamin B_{12} deficiency and liver disease, so a high FIGLU excretion is not specific for the diagnosis.

A deficiency of vitamin B_{12} is investigated by measuring its serum concentration and by microscopical haematological examination of blood and bone marrow. Serum B_{12} can be measured in isolation or as part of a Schilling test to exclude intrinsic factor

Figure 10.44 Molecular model of glutathione reductase. The bound flavin adenine dinucleotide (FAD) coenzymes are shown in red. PDB file 3GRS.

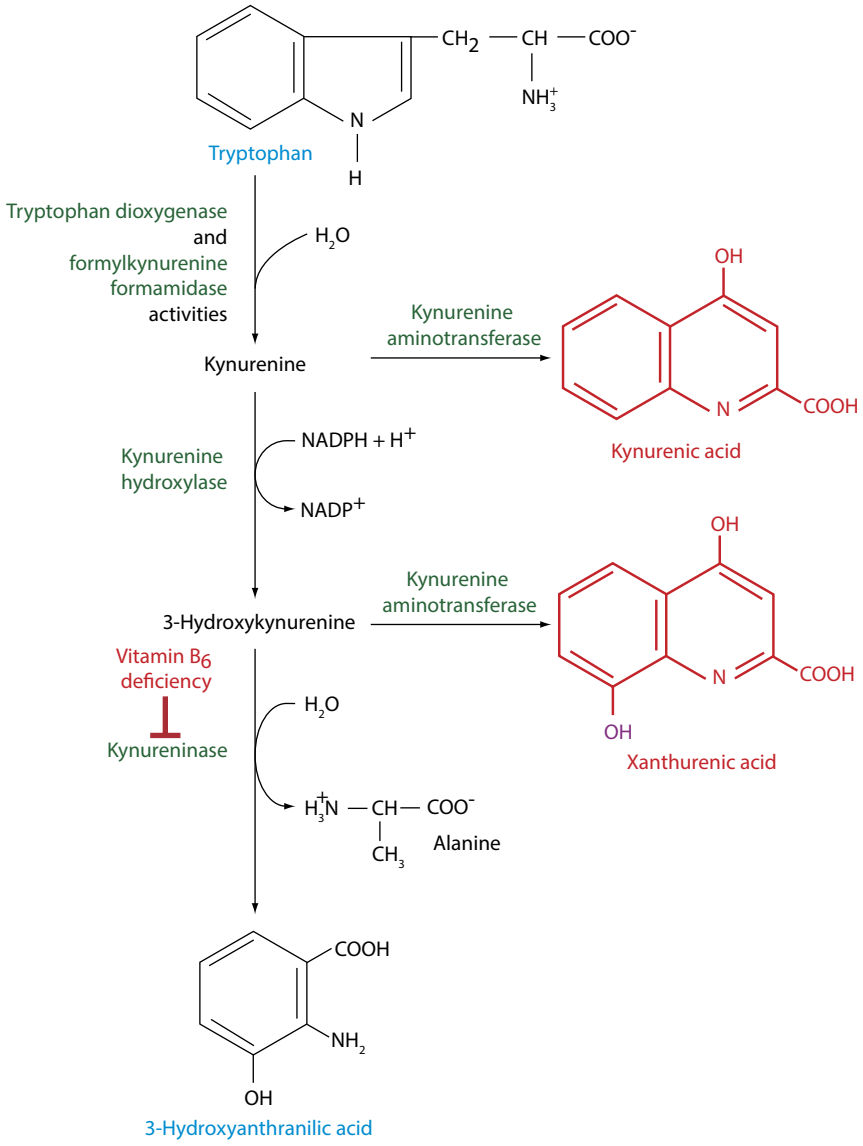


Figure 10.45 The catabolism of xanthurenate, which is the basis of the tryptophan load test to assess vitamin B_6 status.

deficiency (pernicious anaemia; *Chapter 13*). The Schilling test assesses whether vitamin B_{12} is being absorbed appropriately by the body. The amount of vitamin B_{12} excreted in urine over a 24-hour period is determined after giving the patient a known amount of radioactively labelled vitamin B_{12}. If the GIT is able to absorb vitamin B_{12} normally, then up to 25% of the vitamin will be present in the urine. If there is failure in absorption, then little or no vitamin B_{12} is detected in the urine. In the latter case, the test is repeated following an oral dose of intrinsic factor to determine whether the vitamin deficiency is due to a lack of intrinsic factor or GIT problem.

The concentration of vitamin C (ascorbate) in the plasma is a poor indicator of deficiency. Measurements of cellular stores, especially in leukocytes, are more useful. However, vitamin C concentrations in leukocytes should always be accompanied by a differential leukocyte count, given that different types of leukocytes vary in their capacity to accumulate the vitamin. A change in the proportion of polymorphonuclear leukocytes, which become saturated with vitamin C at lower concentrations than other leukocytes, would result in a change in total concentration of vitamin C per 10^6 cells, even if the nutritional status of the vitamin is unchanged.

Fat-soluble vitamins

The concentration of vitamin A in the plasma can be determined, but the values obtained may be misleading, as they only decline when tissue stores become severely depleted. Deficiency can also occur in severe protein deficiency, which decreases the amount of

its carrier protein. In such cases, the concentration of plasma vitamin A would increase once the protein deficiency was corrected.

Clinical investigations of possible vitamin D deficiency involve determining serum calcium and phosphate concentrations and measuring serum alkaline phosphatase activity, since the enzyme is lost from cells during bone catabolism. The metabolite 25-hydroxycholecalciferol in samples of plasma can be measured directly and is a good indicator of vitamin D status in the presence of normal renal function.

Vitamin E (α-tocopherol; see *Figure 10.21*) status can also be assessed by direct measurement of its concentration in the plasma or serum, while vitamin K deficiency is investigated by assessing the prothrombin time of the patient (*Chapter 13*).

Minerals and trace elements

A diagnosis of an overt mineral deficiency or excess can be confirmed by chemical tests that measure its concentration in the serum or, in some cases, urine. These measurements can give an indication of the amounts present in body tissues, although urine values tend to reflect dietary intake rather than the amounts stored in the body. However, clinical/chemical tests for possible mineral and trace element deficiencies must always be used to complement the medical history and physical examination of the patient, since many of their findings may reflect underlying clinical conditions additional to the nutritional status of a patient. It is, therefore, necessary to understand these illnesses and how they influence the findings of physical and laboratory investigations. Measuring the concentrations of trace elements in clinical samples is complex and requires sensitive techniques because of their low values. Samples of plasma are often used, but the determined values may not accurately reflect the concentration of the trace element at its site of action, which may be intracellular. However, for many trace elements, a low plasma concentration is indicative of a deficiency and adequate supplementation needs to be provided, while a high value is an indicator of possible toxic levels.

The value of determining the serum concentrations of sodium, potassium, calcium and phosphate are considered in *Chapter 8*. Investigations of chloride and sulphate are of little clinical relevance.

There is no satisfactory clinical test for chromium deficiency and a diagnosis is usually made following improved glucose tolerance after chromium supplementation.

Copper deficiency is uncommon, except in patients on synthetic oral or intravenous diets. In these patients, plasma copper is reduced to less than 12 μmol/L. Low concentrations of plasma copper may indicate depleted stores but are a poor indicator of short-term copper status. Measuring Cu/Zn superoxide dismutase and cytochrome oxidase activities can also indicate copper status, as the activities of these enzymes are reduced on a low-copper diet.

Serum fluoride concentration can indicate fluoride exposure and may provide information on endemic fluorosis and allow preventative measures to be taken. Some, but not all, studies have reported a direct relationship between serum fluoride and the degree of fluorosis. Measurements of serum iodine may be a useful assessment of thyroid activity in adults. They may also be of use in investigating cretinism in infants, allowing a diagnosis to be made at an earlier age than is possible by other methods. Serum iodine measurements are also of value in assessing iodine toxicity.

Methods for investigating the iron status of a patient are described in *Chapter 13*. Serum iron and total iron-binding capacity can be investigated, but serum ferritin is a better measure of total body iron stores. The amount of magnesium in serum is less than 1% of the total in the body and its measurement is therefore a relatively poor indicator of magnesium status. If hypomagnesaemia is present (*Chapter 8*), then magnesium deficiency is likely but a normal serum value does not exclude a significant deficiency. The determination of molybdenum concentrations in biological fluids is rarely required, which is perhaps fortunate, since the methods used are inadequate due to the low concentrations involved. Selenium deficiency occurs with poor dietary intake and can be detected by its measurement in plasma or whole blood. However, determining erythrocyte activity of the selenium-dependent enzyme glutathione peroxidase (*Figure 10.46*) can indirectly assess selenium status. Zinc concentrations less than 8 μmol/L in the plasma may indicate zinc deficiency, but low values may be associated with hypoalbuminaemia, as most zinc is bound to albumin.

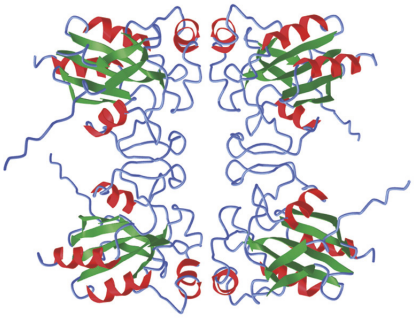


Figure 10.46 Molecular model of glutathione peroxidase. PDB file 2I3Y.

10.5 GENERAL MANAGEMENT OF NUTRITIONAL DISORDERS

Patients who are malnourished or at risk of developing malnutrition require appropriate therapy, which ranges from simple dietary advice to long-term parenteral nutrition.

The dietary needs of the patient must be carefully assessed to provide the appropriate amounts of energy, protein, EFAs, vitamins, minerals and trace elements. Patients receive these diets by oral (*Figure 10.47*), tube and parenteral feeding; the last is most commonly administered by intravenous infusion.

Figure 10.47 An example of a liquid food suitable for oral feeding. ⏎

Oral supplementation should be used wherever possible and the common practice is to encourage consumption of specific foods or supplements that rectify the nutritional disorder in question. In cases where oral feeding is not possible, then liquid food is administered through a nasal tube to the stomach or small intestine. Tube feeding is particularly useful in patients with swallowing difficulties or anorexia. During tube feeding, liquid may be pumped continuously at a constant rate between 75 and 150 mL per hour for 8–24 hours. Liquid foods for tube feeding are available commercially as formulae that meet nearly all the patient needs. In some cases, food is administered as a bolus, that is, infusing a discrete volume of formula through the tube under gravity several times daily. It has the advantages of reducing the cost and giving stable long-term patients more mobility.

Parenteral feeding bypasses the GIT so nutrients are delivered directly into the blood. It is only used when oral or tube feeding have been deemed unsuitable, such as in patients who cannot eat or absorb food from the GIT. Total parenteral nutrition given via the peripheral veins or, in cases of long-term nutrition, through a central venous catheter can provide complete nutrition using preparations containing appropriate amounts of energy, amino acids, fatty acids, vitamins, minerals and trace elements. Like liquid foods for tube feeding, these preparations are available commercially, but they are occasionally prepared to meet the specific needs of an individual patient. Patients on long-term total parenteral nutrition require careful clinical and laboratory monitoring. Indeed, patients often have an increased risk of infection at the venous catheter site so care is necessary. Biochemical changes usually precede any clinical signs of nutrient deficiency and so regular laboratory monitoring is essential for early detection of any micronutrient deficiency.

Patients with PEM cannot immediately accept normal food because there are digestive enzyme deficiencies and often gastroenteritis. Rehydration is a priority and oral solutions (*Figure 10.48*) achieve this in some cases, while intravenous infusions may be necessary in severe cases. Diluted milk with added sugar may be given initially and, as this becomes accepted, the proportion of milk can be gradually increased. The cessation of diarrhoea indicates that the health of the GIT mucosa is improving, and normal foods can be gradually introduced into the diet.

The management of obesity aims to reduce food intake, particularly total energy intake, and to encourage regular exercise. This is often achieved by cutting down on high-energy foods, such as fats and alcohol. Education and psychological support can be helpful in cases of severe obesity. Orlistat and liraglutide used in conjunction with energy-controlled diets help control body weight and so manage obesity. The former is an inhibitor of pancreatic lipase and reduces the digestion and absorption of dietary fats (*Chapter 11*), while liraglutide imparts a feeling of dietary satiety and so reduces appetite

A)
CODE 15-611-10
unicef
United Nations Children's Fund
ORAL REHYDRATION SALTS BP
FOR THE TREATMENT OF DEHYDRATION DUE TO DIARRHOEA
SELS DE REHYDRATATION ORALE BP
POUR LE TRAITEMENT D'ETATS DE DESHYDRATATION LORS DE DIARRHEES
SALES DE REHIDRATACIÓN ORAL BP
PARA EL TRATAMIENTO DE LA DESHIDRATACIÓN CAUSADA POR LA DIARREA
أملاح للماء الفموي
(أملاح للارواء بالفم)
لمعالجة التجفاف الناجم عن الاسهال
Each sachet contains: Glucose anhydrous Food Grade 20.0 g
Sodium chloride BP 98 3.5 g
Trisodium citrate dihydrate BP 98 2.9 g
Potassium chloride BP 98 1.5 g
Net wt.: 27.9 g
Batch No.
Mfg.Date
Exp. Date
B)
Dioralyte Relief
Clinically proven to reduce diarrhoea and dehydration

Figure 10.48 Oral rehydrating solutions. (A) The informative cover of a United Nations-approved sachet containing ingredients that form an oral rehydrating solution appropriate for children when dissolved in a suitable volume of water. (B) A commercially available oral rehydrating solution.

(*Figure 10.49*). Surgery is used in some cases of severe obesity. Jejunoileal bypass surgery for morbid obesity was first performed in 1952. In this process, the end of the jejunum or the beginning of the ileum is removed and the remaining portions joined together. Most of the small intestine is removed, leaving only a short length for digestion and absorption. Although jejuno-ileal bypass surgery results in very good weight loss, severe side effects occur. The technique has been replaced by gastric bypass surgery where the upper part of the stomach is connected to the small intestine about one-third of the way along its length. Thus only the lower part of the stomach and approximately two-thirds of the small intestine are available for digestion and absorption, reducing energy intake and therefore weight, but dieting is still likely to be required. Again, this operation is not without the risk of developing serious postoperative complications. Liposuction may also be used to remove fat from under the skin, although this is typically performed mainly for cosmetic rather than therapeutic reasons.

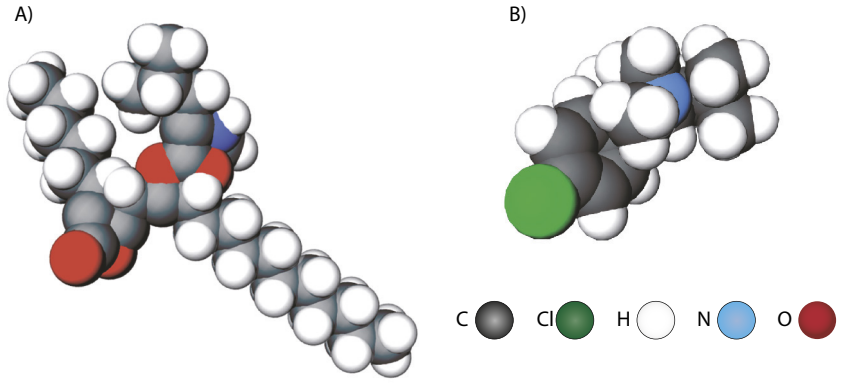


Figure 10.49 Computer drawn models of (A) orlistat and (B) sibutramine.

Eating disorders, such as AN and BN, are difficult to prevent and remain hard to treat. There are seemingly few effective treatments and there is no universally recognised plan of treatment. However, the goals of any therapy must be to establish normal eating patterns and to restore the patient's nutritional status and weight. Usually, a multidisciplinary approach tailored to the individual is used. This will involve specialists in nutrition and mental health in addition to clinicians. Therapy generally involves the family, behaviour modification and nutrition counselling, support groups and the use of antidepressants. Most patients with AN or BN are treated as outpatients, although in severe cases hospitalisation may be necessary. Given that compliance is often problematic, AN and BN are generally considered to be chronic disorders interrupted only by intermittent periods of short-lived remission.

10.6 CASE STUDIES

CASE STUDY 10.1

Andrew, a 51-year-old solicitor, had a health check. He was found to weigh 105 kg and his height was 1.82 m. Andrew is a smoker and does not exercise on a regular basis. He has a family history of heart disease. His blood pressure and plasma lipids were analysed and the results are shown below (reference ranges are shown in parentheses):

Blood pressure	170/100 (120/80 mmHg)
Total cholesterol	6.3 (<6.5 mmol/L, acceptable; <5.2 mmol/L, desirable)
Triacylglycerols	2.7 (0.45–1.80 mmol/L)
HDL cholesterol[a]	0.7 (1.0–1.5 mmol/L)

HDL, high-density lipoprotein.
[a] HDL cholesterol is often colloquially referred to as 'good' cholesterol (see *Chapter 14*).

Questions

(a) Calculate Andrew's BMI.
(b) Is Andrew obese?
(c) What risks are associated with obesity?
(d) What advice would you give Andrew?

CASE STUDY 10.2

Emily is a 17-year-old female. Over the past 18 months, her periods have become increasingly irregular and intermittent and she consulted her doctor about this problem. During the consultation, Emily explained that she was studying for her preuniversity examinations and wishes to study law. Her parents and teachers expect her to do extremely well. She thought she was 'a bit overweight' and was dieting and about two years ago had started exercising more regularly. She is now jogging about two hours a day and has recently taken up aerobic classes. A physical examination showed Emily to be 5′ 4″ tall and weigh 110 pounds. Body temperature, pulse and blood pressure were all normal. A urine Ketostix gave a normal result. Blood tests gave the following results (reference ranges are shown in parentheses):

Glucose	4.3 (4.5–5.6 mmol/L)
Ketone bodies	110 (~70 μmol/L)

Questions

(a) Is Emily underweight?
(b) If so, by how much?
(c) Account for the clinical results.
(d) What treatment or advice would you give Emily?

CASE STUDY 10.3

John, a 74-year-old man residing in a nursing home, was seen by his doctor because he suffered from mental confusion and had difficulty walking because of paraesthesiae and numbness in his legs. There was some general concern that the elderly residents at this nursing home were not being fed appropriately. A blood sample was taken and assessed and an erythrocyte transketolase activity test performed; the results are shown below:

Transketolase activity without added TPP	1.80 mmol/h per 10^9 erythrocytes
Transketolase activity with added TPP	2.21 mmol/h per 10^9 erythrocytes

Questions

(a) Does John suffer from a deficiency of any vitamin?
(b) If so, how should he be treated?

CASE STUDY 10.4

BaoQiang, a three year old, was referred to his local hospital because of ongoing and persistent anaemia. On examination, he appeared pale with oedema in the legs and weighed 11.5 kg (normal weight for a three-year-old boy: 14–15 kg). BaoQiang was mentally listless. It was noticed that his hair had lighter coloured streaks and was easily pluckable. Further examinations revealed an enlarged liver. However, blood electrolytes and tests for kidney function were normal. His serum albumin and total protein were reduced (reference ranges are shown in parentheses):

Serum albumin	15	(32–48 g/L)
Total protein	47	(60–80 g/L)

The family appeared caring but, on questioning, his parents appeared to be financially stressed. Meals consisted largely of small helpings of boiled white rice supplemented by frugal helpings from a charitable food bank, which included some bread, biscuits and potato crisps.

Questions

(a) What is the most likely diagnosis?
(b) Give a plausible explanation for BaoQiang's clinical condition.
(c) How should BaoQiang be managed?

CASE STUDY 10.5

Henry, an eight month old, is failing to thrive, with both weight and height below that expected for a baby of his age. His mother and father are of normal height and there is no family history of malabsorptive conditions. Laboratory investigations to determine the serum concentrations of electrolytes and thyroid-stimulating hormone (TSH), perform a full blood count assessment and evaluate kidney and liver functions showed all were within their reference ranges. However, his serum vitamin D concentration was reduced (reference range given in parentheses):

Vitamin D	15 (50–75 nmol/L)

Questions

(a) If the vitamin D deficiency is allowed to persist, what is the likely clinical outcome?
(b) Why might Henry be vitamin D deficient?
(c) Explain which clinical investigations should be initiated and why they are necessary.

10.7 SUMMARY

The body requires an adequate supply of proteins, carbohydrates, lipids, vitamins, minerals, trace elements and water to provide energy and the raw materials for biosynthesis and to maintain health. The amount of each nutrient required can be expressed as its DRV. Nutritional disorders may arise from deficiencies or, in some cases, an excess of nutrients. Such disorders may be investigated, initially by taking a medical history, by the use of dietary questionnaires and by a physical examination, which may include taking anthropometric measurements. Laboratory tests may be used to confirm suspected diagnoses. Protein–energy malnutrition may lead to marasmus or kwashiorkor, while obesity, an increasing problem in developed countries, predisposes an individual to a variety of disorders including type 2 diabetes and CHD. Anorexia nervosa and BN are complex eating disorders with underlying psychological bases that require treatment, as well as help for the obvious nutritional deficits. The variety of disorders arising from vitamin deficits reflects the range of vitamins required for health and their diverse roles in the body. Hypervitaminoses are uncommon, although some B vitamins along with vitamin A and D toxicities have been observed when excessive amounts of vitamin supplements have been consumed. Disorders associated with mineral deficiencies are rare, although they do occur, for example in the development of thyroid goitre in diets deficient in iodine.

The management of nutritional disorders may require simple dietary advice or, when severe, nutritional supplements. Medical interventions may require tube or parenteral feeding to restore a good nutritional status.

QUESTIONS

1. Calculate the mean energy requirements of:
   (a) a 13-year-old girl weighing 77 kg
   (b) a sedentary adult woman of 40 years of age weighing 62 kg
   (c) a physically active 42-year-old male weighing 73 kg
   (d) an eight-week-old baby weighing 12 lb 7 oz.
   Assume 1 kg = 2.2 lb
   Hint: *Table 10.1* will help.

2. Copper in blood is bound to which protein?
   (a) Thyroglobulin.
   (b) Ferritin.
   (c) Ceruloplasmin.
   (d) Haemoglobin.
   (e) Albumin.

3. Which of the following is NOT a complication of obesity?
   (a) Type 2 diabetes mellitus.
   (b) Vascular disease.
   (c) Hypertension.
   (d) Coeliac disease.
   (e) Cholelithiasis.

4. Arrange the following two lists into their most appropriate pairings.

ATP7A	Xanthine oxidase
Bulimia nervosa	Dental erosion
Copper	Acetyl CoA
Iodide	Dental caries
Fluoride	Wilson disease
Folic acid	Prothrombin time
Leptin	Wernicke–Korsakoff syndrome
Molybdenum	Phytanic acid
Oedema	Osteomalacia
Pantothenic acid	Heart disease
Saturated fats	Haemolytic anaemia
Thiamine	Scurvy
Vitamin B_1	Green leafy vegetables
Vitamin B_6	Thiaminase
Vitamin B_{12}	Tryptophan
Vitamin C	*OBESE/obese*
Vitamin D	Kwashiorkor
Vitamin E	Crohn disease
Vitamin K	Menkes disease
Zinc	Goitre

5. Your adolescent daughter tells you she is going to become a vegan (no consumption of animal products). In light of *Table 10.9* and your knowledge of protein nutrition, what advice would you give her?

Plant food	Deficient in the following amino acids
Cereals:	
maize	K, M, W
rice	H, W
oats	M, W
wheat	K, T, W
Legumes	M, I, W
Nuts	K, T
Oil seeds	K, T
Other vegetable	M, I

Table 10.9 Deficient amino acids in selected vegetable foods.

6. A 72-year-old man who lives by himself was admitted to hospital following an accident. On admission, it was noted that he appeared emaciated, anaemic and had widespread petechial haemorrhages. He also presented with gum disease. What types of nutritional deficiencies might be present?

7. *Figure 10.41* shows the arm circumference and skinfold thickness being measured on two male patients, A and B, respectively. *Table 10.10* gives additional anthropometric data for them.

Metric	A	B
Height	1.83 m	1.80 m
	6′ 0″	5′ 10″
Weight	77.5 kg	120 kg
	171 lb	266 lb
Triceps skinfold thickness (mm)	12	15
Arm circumference (mm)	270	320
WHtR	0.44	0.66

Table 10.10 Selected anthropomorphic data of male patients A and B. ⏎

What can be concluded from these data about the respective health risks to A and B?

FURTHER READING

Amrein K *et al*. (2020) Vitamin D deficiency 2.0: an update on the current status worldwide. *Eur J Clin Nutr* **74**:1498–1513.

Azzi A (2018) Many tocopherols, one vitamin E. *Mol Aspects Med* **61**:92–103.

Bender DA and Cunningham SMC (2021) *Introduction to Nutrition and Metabolism*, 6th edition. Taylor & Francis.

Beulens JW *et al*. (2013) The role of menaquinones (vitamin K_2) in human health. *Br J Nutr* **110**:1357–1368.

Bhutta ZA *et al*. (2017). Severe childhood malnutrition. *Nat Rev Dis Primers* **3**:17067.

Castillo M and Weiselberg E (2017) Bulimia nervosa / purging disorder. *Curr Probl Pediatr Adolesc Health Care* **47**:85–94.

Diab L and Krebs NF (2018) Vitamin excess and deficiency. *Pediatr Rev* **39**:161–179.

Galli F *et al*. (2017) Vitamin E: emerging aspects and new directions. *Free Radic Biol Med* **102**:16–36.

Kiani AK *et al*. (2022) Main nutritional deficiencies. *J Prev Med Hyg* **63**(suppl. 3):E93–E101.

Lykkesfeldt J (2020) On the effects of vitamin C intake on human health: how to (mis)interpret the clinical evidence. *Redox Biol* **34**:101532. doi: 10.1016/j.redox.2020.101532.

Neale J and Hudson LD (2020) Anorexia nervosa in adolescents. *Br J Hosp Med (Lond)* **81**:1–8.

Oadayatty SJ and Levine M (2016) Vitamin C: the known and the unknown and Goldilocks. *Oral Dis* **22**:463–493.

Piché ME *et al*. (2020) Obesity phenotypes, diabetes, and cardiovascular diseases. *Circ Res* **26**:1477–1500.

Pilz S *et al*. (2019) Vitamin D testing and treatment: a narrative review of current evidence. *Endocr Connect* **8**:R27–R43.

Schwalfenberg GK (2017) Vitamins K_1 and K_2: the emerging group of vitamins required for human health. *J Nutr Metab* **2017**:6254836. doi: 10.1155/2017/6254836.

Smith AD and Refsum H (2016) Homocysteine, B vitamins, and cognitive impairment. *Ann Rev Nutr* **36**:211–239.

Taubes G (2001) Nutrition. The soft science of dietary fat. *Science* **291**:2536–2545.

Ten Cate JM (2004) Fluorides in caries prevention and control: empiricism or science. *Caries Res* **38**:254–257.

Vermeer C and Schurgers LJ (2000) A comprehensive review of vitamin K and vitamin K antagonists. *Hematol Oncol Clin North Am* **14**:339–345.

Useful website

https://www.food.gov.uk

Taylor & Francis
Taylor & Francis Group
http://taylorandfrancis.com

11

Disorders of the gastrointestinal tract, liver, gallbladder and pancreas

OBJECTIVES

After studying this chapter you should be able to:

- outline the structures of the gastrointestinal tract and its accessory organs
- describe how foods are digested
- outline how nutrients are absorbed by the gastrointestinal tract
- review the role of the small intestine in homeostasis
- describe some common gastrointestinal tract disorders
- describe some common pancreatic, liver and gallbladder disorders

11.1 INTRODUCTION

The gastrointestinal tract (GIT) is essentially a long tube, extending from the mouth to the anus, with several specialised regions. Most foods (*Chapter 10*) must be digested before their nutrients can be assimilated. Digestion involves the hydrolysis of oligosaccharides, polysaccharides, peptides, proteins, nucleic acids and lipids in enzyme-catalysed reactions. The products of hydrolysis, such as monosaccharides, amino acids, nucleosides and nucleotides, free fatty acids and monoacylglycerols, can then be absorbed. Some nutrients, vitamins and minerals, for instance, are absorbed without the need for digestion. The digestion of food and absorption of the released nutrients are the major functions of the GIT, together with its accessory organs, the pancreas, liver and gallbladder. The **peritoneum** lines the abdominal cavity and covers the major organs. Folds of the peritoneum, called the **mesentery**, and the **mesocolon** hold the intestines in place. The GIT also has significant roles in the homeostasis of nutrients and in protecting the body against pathogens (*Chapter 2*) ingested with food. Digestion and absorption can, of course, only take place after ingestion of food, unlike the homeostatic and protective functions, which operate continuously.

11.2 GASTROINTESTINAL TRACT

The GIT may be described as a tubular system with distinctive regions that begin with the mouth, leading to the pharynx, oesophagus, stomach, and small and large intestines and terminating at the anus (*Figure 11.1*). It is associated with several accessory digestive organs, such as the pancreas, liver and gallbladder. The walls of the GIT have a fairly common substructure (*Figure 11.2*) and its interior is called the **lumen**.

11.3 ACCESSORY ORGANS OF THE GASTROINTESTINAL TRACT

The major accessory organs of the GIT are the liver, gallbladder and pancreas (*Figure 11.3*).

DOI: 10.1201/9781003296836-11

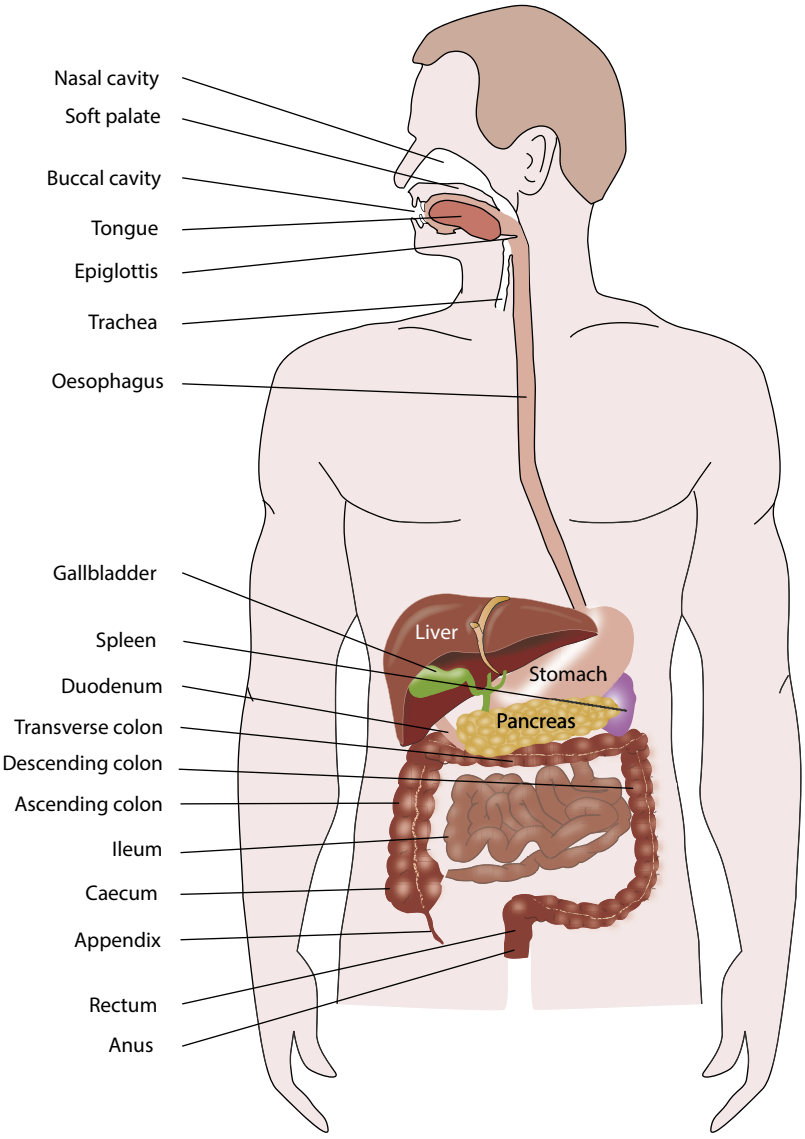


Figure 11.1 Outline of the structure of the human gastrointestinal tract.

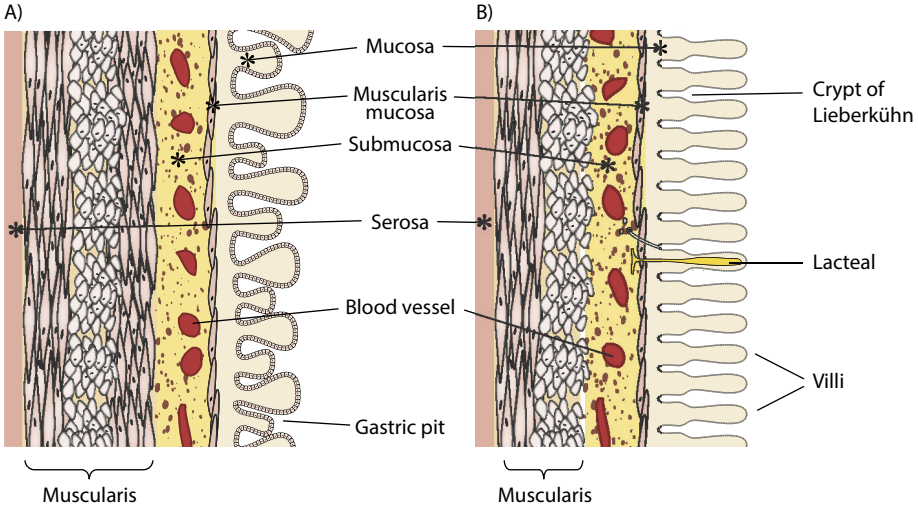


Figure 11.2 Outline of the tissue organisation of the walls of (A) the stomach and (B) the small intestine.

Margin Note 11.1
Flavour and taste

Flavour is a complicated amalgam of sensory perceptions involving the taste, smell and texture of foods, previous exposure to particular foods, especially in childhood, and even the noises associated with chewing some foods. In general, much of the everyday sensation of flavour is derived largely from olfactory stimulation. **Anosmia**, the loss of the sense of smell, for example due to infection with a cold, can severely reduce the sense of taste.

Six tastes can be detected by different types of sensory taste buds distributed over the tongue. These are bitter, salt, sour, sweet, fats (fatty acids) and savoury or umami. The first four are self-explanatory. The perception of fats is more strictly the detection of small amounts of fatty acids hydrolysed from lipids by the action of lingual lipase. In 1907, Kikunae Ikeda (1864–1936), working at Tokyo Imperial University, investigated the source of a flavour well recognised in many Asian countries called umami. This most easily translates in English to savoury or meaty. He was successful in extracting crystals of the amino acid glutamate from a type of seaweed called kombu used in traditional Japanese cuisine. Many foods that have an umami flavour are protein rich, for example meat, seafoods and aged cheeses, and contain large amounts of glutamate. Monosodium glutamate (MSG) was developed as a food additive and is probably the most familiar example of an additive that imparts a savoury or umami taste.

The evolution of specific taste receptors is most easily explained by the advantages gained in detecting foods rich in essential nutrients (salty, sweet, fats and savoury) and in avoiding ones that contain toxins (bitter and sour).

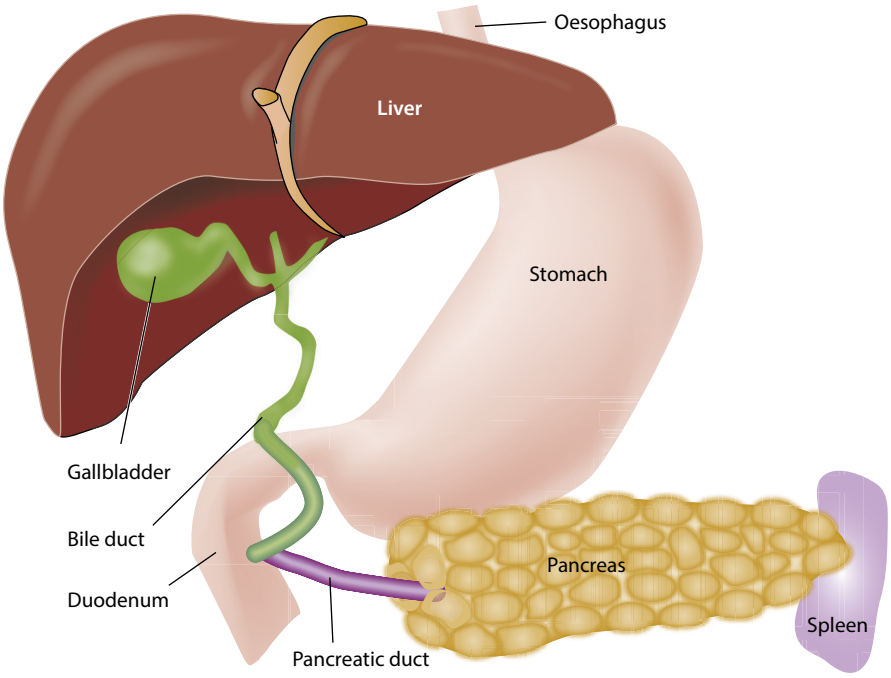


Figure 11.3 The major accessory organs of the gastrointestinal tract.

Liver and gallbladder

The liver is the largest organ in the body, weighing about 1.5 kg. It has two major lobes, each divided into functional units called lobules. Lobules consist of cuboidal cells arranged around a central vein into which the blood drains. **Sinusoids** occur between the plates of hepatic cells. The sinusoids are supplied by the portal vein and hepatic artery and lined with phagocytic Kupffer cells (see *Figure 12.5*) that belong to the mononuclear phagocytic system (formerly reticuloendothelial system). Their main function is to engulf bacteria and other foreign particles in blood. Thus, there is a dual blood supply to the liver, with blood coming from the digestive tract and spleen through the portal vein and from the aorta in the hepatic artery (see *Figure 14.2*). About a third of incoming blood is arterial and supplies oxygen and nutrients; the remainder is venous from the portal vein. The large reserve capacity of the liver means that it needs only 10–20% of its tissues to function to sustain life. It also has a remarkable ability to regenerate itself after its tissue has been removed or destroyed by disease. Complete destruction or removal of the liver results in death within 10 hours; hence, liver disease with loss of function is a serious clinical matter (*Chapter 12*).

The liver has numerous functions. It acts as an interface between the GIT and the rest of the body tissues. Given that the hepatic portal vein carries blood directly to the liver from the GIT, it can regulate the posthepatic blood concentrations of many of the nutrients absorbed by the GIT. Similarly, it also regulates the concentrations of numerous biomolecules produced by the body, for example steroid hormones, and deals with many toxins, such as drugs, pesticides and carcinogens, rendering them less harmful and more soluble to facilitate their excretion (*Chapter 12*). The liver also produces many of the plasma proteins including albumin and the clotting factors (*Chapter 13*). Bile produced by hepatocytes is secreted into bile canaliculi and eventually drains into the bile duct. About 700–1200 mL of bile is produced daily and stored and concentrated in a hollow organ called the gallbladder (*Figure 11.3*) prior to its release in the small intestine. Bile does not contain enzymes, but it does contain bile salts, for example glycocholate and taurocholate (*Figure 11.4*). These are detergents that aid in the digestion of lipids by emulsifying them to form water-soluble complexes.

The spleen (see *Figure 11.3*) destroys dysfunctional erythrocytes and releases haemoglobin. This is then degraded to haem and globin (*Chapter 13*); while the latter is hydrolysed to amino acids, the haem is converted to bilirubin (*Figure 11.5*), which is catabolised, forming bile pigments. A typical adult produces approximately 450 µmol of bilirubin daily. Unconjugated bilirubin is insoluble in water and is transported in the plasma bound to serum albumin to be taken up by hepatocytes. Here, it is conjugated with glucuronic acid by UDP-glucuronyltransferase activity to form mono- and diglucuronides in

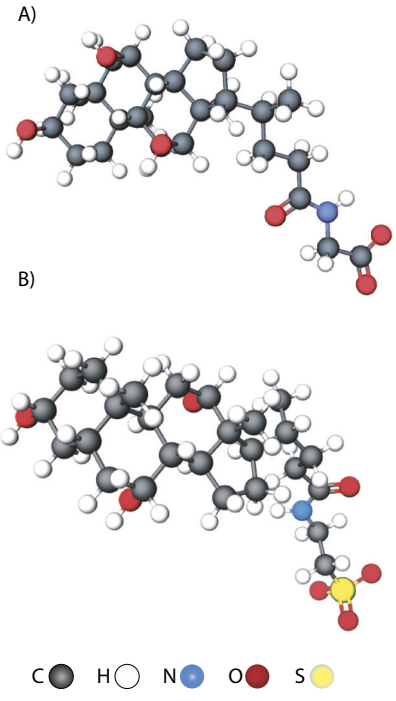


Figure 11.4 Computer drawn models of (A) glycocholate and (B) taurocholate.

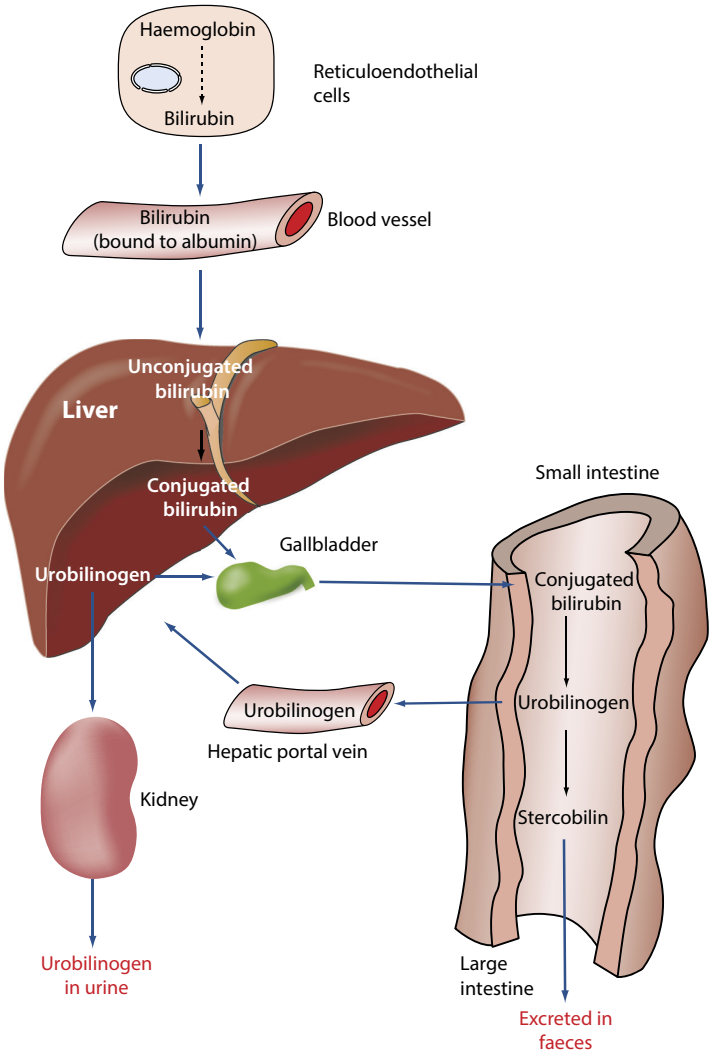


Figure 11.5 Overview of the production and excretion of bilirubin following the catabolism of haem. Transport movements are indicated by blue arrows.

a manner resembling the detoxification reactions described in *Chapter 12*. Conjugated bilirubin is much more water soluble than its free, unconjugated, form and is secreted into the bile duct. In the small intestine, conjugated bilirubin is a substrate for bacteria, which convert it to urobilinogen and stercobilin. Most stercobilin is excreted in faeces, although some urobilinogen is absorbed and taken to the liver in the hepatic portal vein and re-excreted in bile or by the kidneys (*Figure 11.5*).

Pancreas

The pancreas is found near the stomach and small intestine (see *Figures 11.1* and *11.3*) and functions as both an exocrine and an endocrine gland. In its exocrine role, it produces pancreatic juice that is released into the first part of the small intestine, the duodenum, through the pancreatic duct. This secretion contains precursors of digestive enzymes that are activated and function in the GIT, and also increases the pH of GIT contents. This latter function is essential, since the material leaving the stomach is acidic. The pancreas produces about 1200–1500 mL of pancreatic juice daily. Within the tissues of the pancreas are groups of cells called the islets of Langerhans. The cells in the islets have endocrine functions and produce and secrete the peptide hormones glucagon and insulin that regulate the concentration of glucose in the blood (*Chapter 7*).

11.4 PERITONEUM

The GIT and accessory organs are bound to each other and to the inner wall of the abdomen by the peritoneum. This is a strong, colourless membrane with a smooth surface

that consists of two parts: the parietal peritoneum, which lines the abdominal cavity, and the visceral peritoneum that covers most of the organs in the abdomen (*Figure 11.6*). The thin space between the two parts, called the peritoneal cavity, is filled with **serous fluid**. In males, the peritoneum forms a closed sac, but in females it is continuous with the mucous membrane of the uterine tubes. The mesentery is a fan-shaped portion of peritoneum that extends from the posterior abdominal wall and wraps around the small intestine and returns to its origin forming a double layer (*Figure 11.6*). The mesentery contains blood and lymphatic vessels and the nerves that supply the small intestine. Another peritoneal fold, the mesocolon (*Figure 11.6*), has similar functions with respect to the large intestine. The mesentery and mesocolon loosely anchor the intestines in place, which allows them considerable movement as muscular activities mix and move the contents along the GIT.

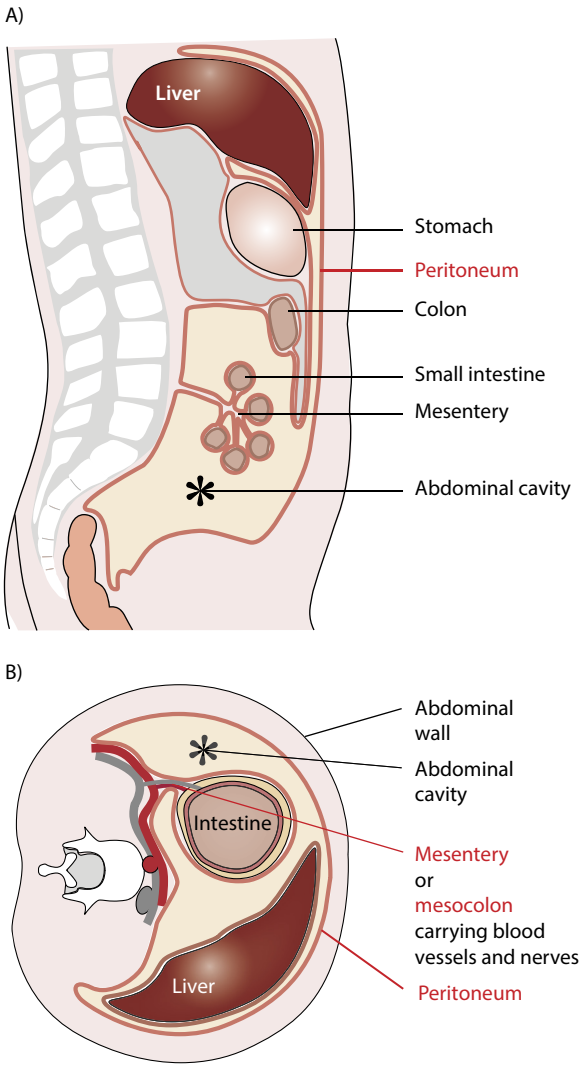


Figure 11.6 Schematics showing (A) the peritoneum and (B) a transverse view of the mesentery and mesocolon.

11.5 DIGESTION

Digestion is the hydrolytic breakdown of food macromolecules and compound lipids to smaller products that can be absorbed. The hydrolytic reactions are catalysed by a variety of enzymes: proteases that digest (hydrolyse) proteins, carbohydrases that digest carbohydrates, nucleases that degrade DNA and RNA, and lipases that catalyse the hydrolysis of lipids. Digestion occurs in the mouth (to a small extent), stomach and small intestine, and most absorption of nutrients occurs in the small intestine and that of water in the large intestine.

BOX 11.1 PERITONITIS

Peritonitis is an inflammation of the peritoneum that lines the abdominal cavity and covers the surfaces of abdominal organs. The condition is marked by exudations into the peritoneum of serum, fibrin, cells and pus. Primary peritonitis is caused by the spread of an infection from the blood and lymph nodes to the peritoneum and accounts for less than 1% of all cases. Secondary peritonitis is the most common type and occurs when bacteria enter the peritoneum from the GIT or biliary tract. The symptoms of peritonitis include vomiting, swelling of the abdomen, severe abdominal pain and tenderness, weight loss, constipation and moderate fever. Its major causes are perforations of the GIT wall, which allows seepage of the luminal contents into the abdominal cavity. These can arise from a ruptured appendix or from perforations of the stomach, including damage from ulcers, and gallbladder. Pelvic inflammatory disease in sexually active women is also a common cause. Peritonitis can also develop after surgery when bacteria have been allowed to enter the abdomen during an operation.

Peritonitis must be rapidly diagnosed and treated because complications can occur rapidly. Diagnosis relies on taking a medical history and a physical examination, particularly to investigate any abdominal swelling and tenderness. Diagnostic tests include using X-rays or a computed tomography (CT) scan (*Chapters 14* and *18*) of the abdomen to confirm the presence of fluid, accumulation of pus or infected organs in the abdomen. Microbiological tests on samples of blood or abdominal fluid can identify the causative microorganism. Peritonitis is frequently life threatening, and acute peritonitis is a medical emergency. The outlook for untreated patients is poor.

Specific treatments for peritonitis depend on age, health, medical history and the severity of the condition. Treatment is generally aimed at rectifying the underlying condition. Antibiotics are given immediately once peritonitis has been diagnosed. A nasal tube may be inserted into the stomach or intestine to drain fluid and gas. Peritoneal lavage, where large amounts of fluid are injected into the peritoneum to wash out the infective microorganisms causing the condition, may be desirable. Intravenous fluids are also given to replace lost body fluids. Morphine (see *Figure 12.2*) may be prescribed in hospitalised patients to reduce pain. Emergency exploratory surgery may be necessary, especially in cases that involve appendicitis, a perforated peptic ulcer or diverticulitis.

Mouth

In the mouth, teeth (*Box 11.2*) break the food into smaller portions increasing the surface area upon which digestive enzymes can act. Three pairs of salivary glands secrete saliva into the mouth: the submandibular located under the jaws, the sublingual sited under the tongue and the parotid situated in front of the ears (see *Figure 11.10*). The saliva contains amylase, lingual lipase, ions (for example Na^+ and Cl^-), water and mucus (see *Margin Note 4.1* and *Box 11.3*). The water helps to dissolve nutrients, while the mucus acts as a lubricant to aid swallowing and the passage of food through the GIT. Chewing and the actions of the tongue mix the food with the saliva. Salivary amylase begins the digestion of carbohydrates (see *Figure 11.11*), although little occurs in the mouth, since the food is not retained there for any length of time. Lingual lipase is secreted by lingual serous glands. Again, little digestion occurs in the mouth but it has been suggested that the fatty acids released by its action are perceived by specific taste receptors allowing fatty foods, and therefore energy-rich nutrients, to be detected.

BOX 11.2 TEETH ⏎

The 20 primary or deciduous (often called baby or milk) teeth erupt from the gums or gingiva at 6–36 months of age. These are gradually lost from the age of five years and replaced by a set of 32 permanent (adult) teeth (*Figure 11.7*).

Human teeth have a general structure comprising three regions (*Figure 11.8*):

1. crown
2. neck
3. root

The outer layer of the crown is formed of an enamel consisting of 96% prisms of hydroxyapatite and only 4% organic matter and water, making it the hardest substance in the body. The enamel covers a layer of hard dentine, comprising 70% hydroxyapatite, 20% collagen and 10% water. This surrounds the pulp cavity containing nerves and lymphatic and blood vessels. The roots of teeth are embedded in alveoli or sockets of the alveolar ridges of the mandible and maxilla. They are

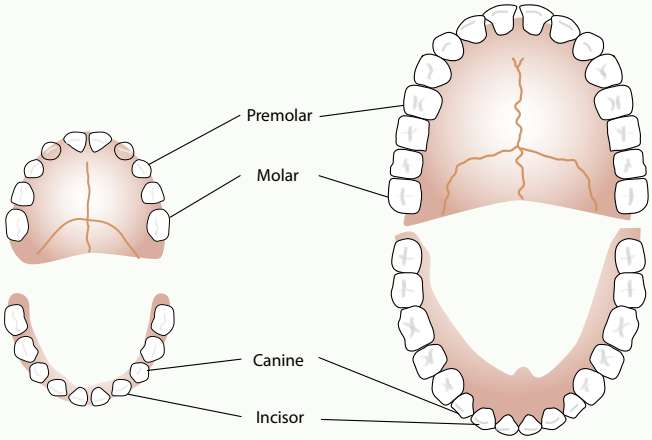


Figure 11.7 The number and arrangement of deciduous (baby or milk) and permanent (adult) teeth. ⏎

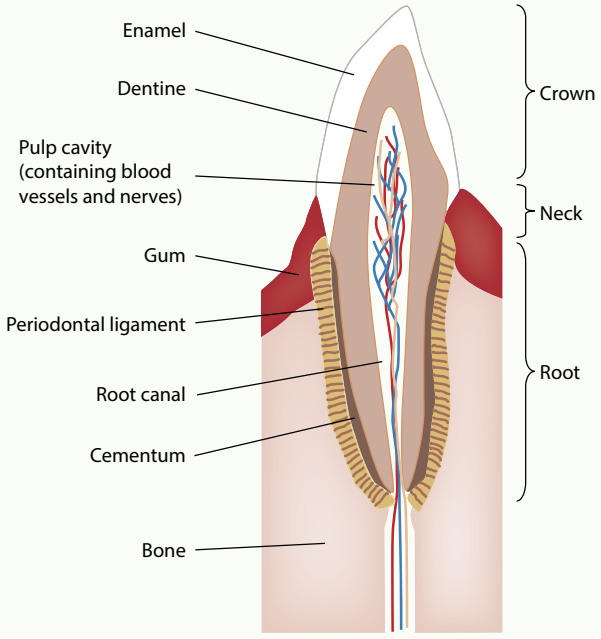


Figure 11.8 The generalised structure of a tooth.

held in place by a thin layer of specialised bone referred to as cementum. This, in turn, is joined to a fibrous connective layer called the periodontal membrane or ligament that allows restricted tooth movements during chewing.

Humans are heterodonts, with four types of specialised teeth adapted for different functions (*Figure 11.9*). Incisors are chisel-like and cut food. The canines are much reduced compared with many mammals but have grip and cutting functions, while the broad, irregular surfaces of premolars and molars function in grinding food. In short, they are a toolset that meets the needs of an omnivore. Teeth reduce food to smaller pieces that have a relatively large surface area on which digestive enzymes can act, and which can be easily swallowed.

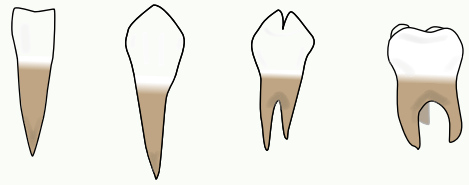


Figure 11.9 Human heterodont teeth.

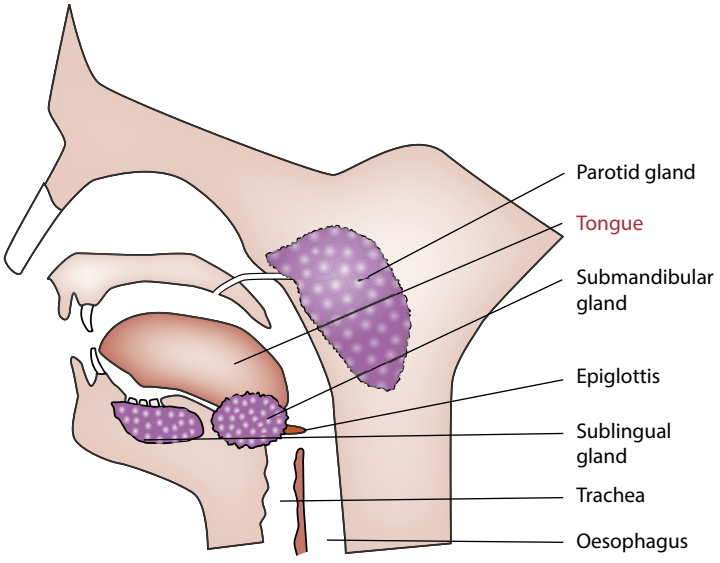


Figure 11.10 The submandibular, parotid and sublingual salivary glands.

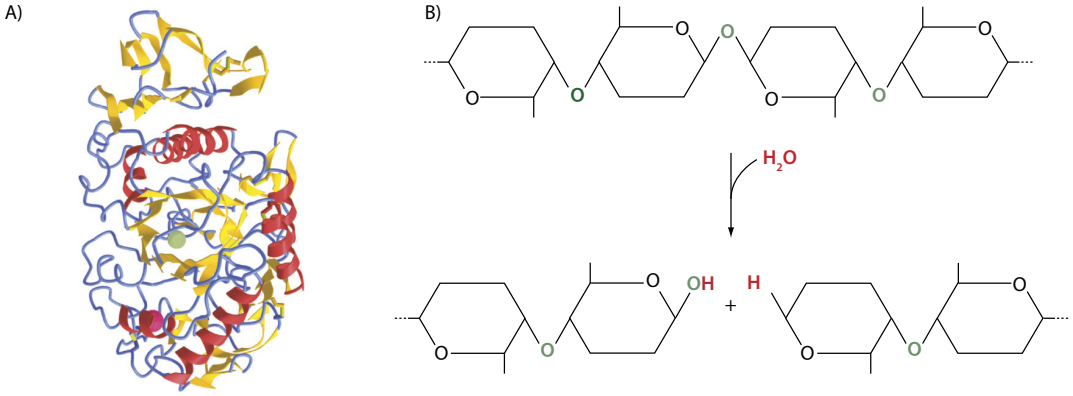


Figure 11.11 (A) Molecular model of a salivary amylase with its bound Ca^{2+} and Cl^- depicted as red and green spheres, respectively. PDB file 1SMD. (B) Schematic outline of the digestion (hydrolysis) of polysaccharide carbohydrates.

In swallowing, the soft palate is elevated, which seals off the nasal cavity and pushes against the back wall of the pharynx. This, in turn, triggers an automatic reflex action in muscles that raises the larynx, pushing its opening, the glottis, against a flap of tissue called the epiglottis. These actions prevent the ball of food or *bolus* entering the trachea of the respiratory system and ensure food is expelled from the mouth into the oesophagus. Rhythmic waves of contraction of smooth muscle tissue in the walls of the GIT called **peristalsis** force food along the GIT. Peristalsis in the oesophagus ensures that food is propelled to the stomach, even if the person is upside down.

Stomach

The stomach is a sac-like region of the GIT with a highly folded inner surface (*Figure 11.12*) allowing it to expand up to eightfold following a meal from its empty volume of about 50 mL. Muscular activities of the stomach wall mix the food with gastric juice secreted by gastric glands located in the mucosal lining of the stomach. Gastric glands produce about 1–2 L of gastric juice per day from three types of secretory cells:

1. parietal or oxyntic cells that produce hydrochloric acid (HCl)
2. peptic or chief cells, which secrete pepsinogen (the precursor of pepsin) and gastric lipase
3. mucous cells that secrete mucus, which protects the mucosa lining the stomach from the corrosive action of the HCl

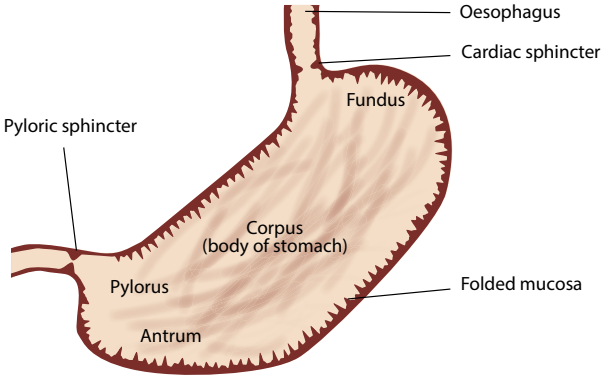


Figure 11.12 The major regions of the stomach.

BOX 11.3 MUCUS AND MUCINS

In the gastrointestinal tract (GIT), mucus occurs as an adherent layer overlayered with more loosely attached mucosal material that together separate the lumen from the underlying endothelium. This layer lubricates the passage of food through the system and protects the endothelial lining from autodigestion, toxins and microorganisms. It is thickest in the colon, where the bacterial load is especially high. In the stomach, mucosal HCO_3^- also offers protection against HCl.

The major component, up to 95%, of mucus is water; the remainder includes glycoproteins, proteins and lipids. The major components responsible for binding such large quantities of water and imparting its viscous, gel-like properties are glycoproteins of M_r up to 20×10^6, called **mucins** (*Figure 11.13*). Mucins are extensively glycosylated, consisting of up to 80% carbohydrate, which renders them resistant to digestive enzymes, allowing them to pass along the GIT and be largely lost in the faeces. Heavily glycosylated membrane mucins contribute to the dense protective glycocalyx of GIT cells (see *Figure 11.18*). Several mucus proteins, such as lysozyme, immunoglobulins and defensins, have defensive functions.

The extended conformations of the mucins means that the carbohydrates 'organise' the water molecules such that the mucosal barrier is not absolute but forms a permeable layer. This allows the passage of biomolecules, secretions, food and nutrients, and respiratory gases between the GIT lumen and its lining of cells.

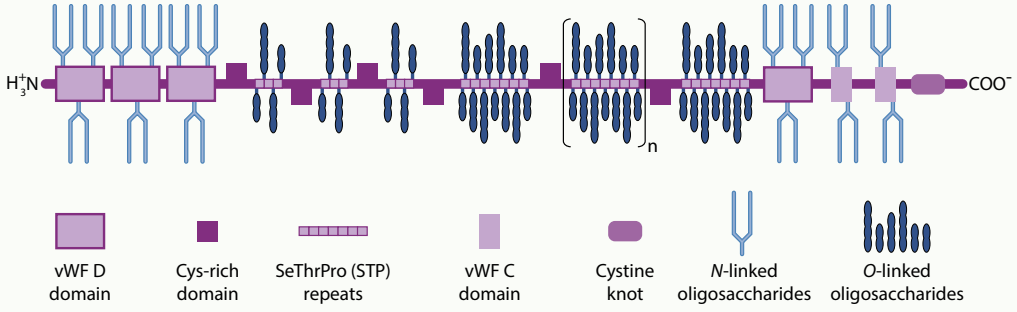


Figure 11.13 Schematic showing the structure of a mucin molecule. Adapted from Bansil R and Turner S (2005), *Curr Opin Colloid Interface Sci* 11:164–170. With permission from Elsevier.

The activity of salivary amylase continues within the bolus and is possibly only inactivated when the bolus contents are completely mixed with the acidic gastric juice. The pH optimum of lingual lipase is between 3.5 and 6.0 and is activated in the upper portion of the stomach. Gastric lipase (*Figure 11.14*) has an acid optimum. About 10–30% of dietary fat may be hydrolysed in the stomach. However, lipase activity in the stomach is highest against triacylglycerols with short or medium-length fatty acid residues, which are found in milk, and so may be of most importance in babies.

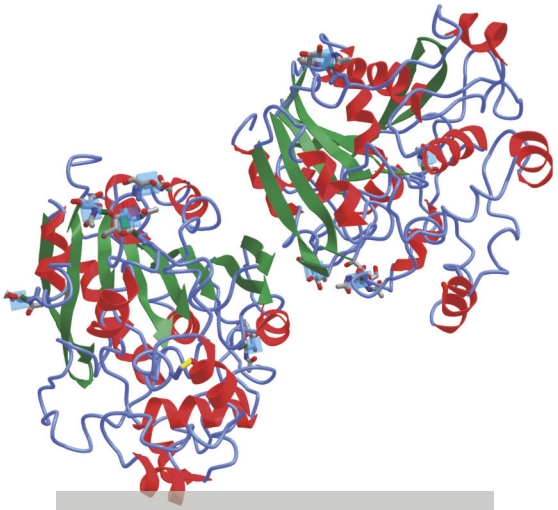


Figure 11.14 Molecular model of gastric lipase. The grey block shows the approximated position of the plasma membrane. PDB file 1HLG.

Margin Note 11.2
Vitamin B_{12} (cobalamin) and gastric intrinsic factor

The absorption of vitamin B_{12} (cobalamin) is dependent upon gastric intrinsic factor (GIF). Thus, any condition that decreases the secretion of GIF, for example atrophic gastritis (*Section 11.10*); interferes with digestion of the binding proteins, such as pancreatic exocrine insufficiency; or decreases the absorption of the GIF–vitamin B_{12} complex, for instance inflammatory bowel disease, can cause pernicious anaemia (*Chapter 13*).

Figure 11.15 Molecular model of a pepsin molecule. PDB file 3UTL.

Pepsinogen is a weakly active protease. It has an acid optimum pH, hence the need for HCl secretion in the stomach. In these conditions, about pH 1–2, protein molecules in the food are denatured, making them more susceptible to digestion. Also, at this pH, pepsinogen molecules act on one another at specific sites to produce the fully active protease pepsin (*Figure 11.15*), which begins the digestion of the denatured proteins to form shorter polypeptides and peptides. The denaturation and digestion of proteins reduce the chance of their absorption and prevent them being immunogenic (*Chapter 4*). The watery mixture of gastric juice and partially digested food is called **chyme**. Vitamin B_{12} (cobalamin) is released from dietary proteins by the action of pepsin and binds to haptocorrin (*Figure 11.16*), one of two binding proteins, the other being gastric intrinsic factor (GIF), which is also present in gastric juice (see *Margin Note 11.2*). The affinity of haptocorrin for B_{12} is increased in acid conditions and is greater than that of GIF. However, when chyme enters the small intestine, vitamin B_{12} becomes bound to GIF and the complex is absorbed by cells of the mucosal lining (*Section 11.6*).

The acid environment of the stomach kills most of the bacteria that are ingested with the food, but some survive and enter the small intestine in the chyme. The pyloric sphincter controls the passage of chyme from the stomach to the small intestine, allowing only relatively small amounts of chyme through at any one time. This control is necessary because the digestive activity within the small intestine is time consuming and because its capacity is limited. The activities of the stomach and small intestine are coordinated by the nervous and endocrine systems.

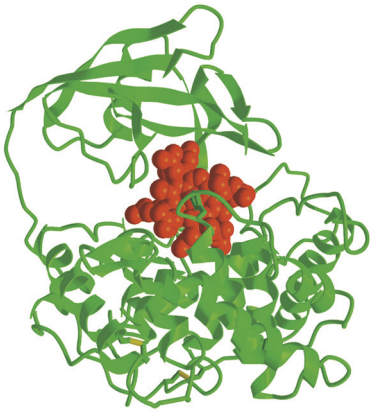


Figure 11.16 Molecular model of haptocorrin in green, with its bound vitamin B_{12} (cobalamin) highlighted in red. PDB file 2BB5.

Small intestine

The final stages of digestion and the absorption of its products occur in the small intestine. Contractions in the small intestine help to break food up, mix it with digestive juices

BOX 11.4 VOMITING AND CYCLIC VOMITING SYNDROME

Vomiting or emesis is the expulsion of food from the stomach (and sometimes duodenum) through the oesophagus and mouth. It is usually experienced as the final of three events: nausea, retching and vomition. Nausea is an unpleasant psychic experience associated with decreased gastric motility and increased muscle tone in the small intestine. Additionally, there is often reverse peristalsis in the proximal small intestine. Retching or dry heaves are the spasmodic respiratory movements conducted with a closed glottis. While this is occurring, the antrum of the stomach contracts and the fundus and cardia relax. Vomition is when the contents of the stomach and sometimes small intestine are propelled up to and out of the mouth.

Vomiting can be initiated by a variety of stimuli, including various chemical irritants (emetics) and poisons, distension of the stomach, unpleasant sights and smells, dizziness, anaesthetics, a number of drugs, a variety of illnesses and conditions (for example gastroenteritis and brain tumours, respectively), pathogens and parasites, and hormonal changes associated with pregnancy. All result in nervous impulses being sent to a vomiting centre in the medulla oblongata (see *Figure 7.4*), which responds by sending a motor impulse to muscles of the upper GIT, diaphragm and abdominal muscles. The resulting contractions squeeze the stomach between the diaphragm and abdominal muscles and expel its contents. A simple episode of vomiting rarely causes problems but may on occasion have more serious consequences, such as aspiration pneumonia. However, severe or repetitive vomiting results in disturbances to acid–base balance, electrolyte depletion and dehydration. In such cases, the underlying cause must be rapidly identified and appropriate therapy initiated. In many cases, antiemetic drugs (for example prochlorperazine, metoclopramide, domperidone and cyclizine) must be administered to suppress vomiting and reduce its sequelae.

Cyclic vomiting syndrome (CVS) or abdominal migraine is a disorder of the GIT characterised by recurrent, severe and prolonged attacks of nausea, vomiting and abdominal pain that was first described by Samuel Gee (1839–1911) in 1882. Its incidence and prevalence are also uncertain, but some evidence suggests that 1 in 50 children in the USA may be affected. Cyclic vomiting syndrome usually occurs in children 2–16 years old, most commonly between the ages of 3 and 7 years, but can also occur in adults. The cause of CVS is unknown.

Cyclic vomiting syndrome occurs in four distinct phases:

1. prodrome
2. episode
3. recovery
4. interval

The prodrome or first phase is often accompanied by abdominal pain and signals that an episode of nausea and vomiting is about to begin. This can last a few minutes to several hours. However, sometimes prodromes do not occur and patients enter an episode directly. Episodes consist of nausea followed by severe vomiting, which usually begin at night or first thing in the morning. Abdominal pain, dizziness, headaches, photosensitivity, fever and sometimes diarrhoea may also present. Vomiting can be as frequent as 6–12 times an hour during the worst of the episode and can continue for one to five days, during which the patient appears pale, listless and exhausted, often to the point of near unconsciousness. Episodes are triggered by specific events or conditions, most commonly infections such as colds and influenza, but also emotional stress, excitement, allergies, eating certain foods (for example chocolate or cheese), overeating, excessively hot weather and physical exhaustion. Motion sickness and menstruation can also trigger episodes. A recovery phase begins when the nausea and vomiting stop. Generally, appetite and energy return, but the time required for this varies considerably. An interval is the symptomless phase between episodes.

The severity of vomiting in CVS is a risk factor for several complications, the most obvious being dehydration and electrolyte imbalance, which are described in *Chapters 8* and *9*. Other complications include peptic oesophagus caused by stomach acid in the vomit; haematemesis, which is blood in the vomit from the damaged oesophagus; and a Mallory–Weiss tear of the lower end of the oesophagus or the stomach, which may bruise from violent and prolonged vomiting or retching. Finally, tooth corrosion can be caused by stomach acid in a manner similar to that seen in bulimics (*Chapter 10*).

Cyclic vomiting syndrome is difficult to diagnose and many patients are initially misdiagnosed. It is usually identified from the general symptoms and medical history and by excluding more common diseases or disorders that can also cause nausea and vomiting. Diagnosis is time consuming because a repeating pattern of vomiting must be established. There is no cure for CVS. Avoiding known triggers of an episode is an obvious remedy, but they cannot always be avoided. Ibuprofen taken in the prodrome phase may help prevent or alleviate an episode. Once an episode begins, treatment is supportive, with bed rest and sleep in a dark, quiet room. Other medications that may be helpful are H^+ pump inhibitors, for example lansoprazole or omeprazole, which decrease the amount of stomach HCl produced. Severe nausea and vomiting may require hospitalisation and intravenous fluids to prevent dehydration, since drinking water normally promotes more vomiting, although it does dilute the acid in the vomit, making the episode less painful. Sedatives may help if the nausea continues. Drinking water and replacing lost electrolytes is essential in the recovery phase.

Cyclic vomiting syndrome and migraine show a number of similarities and the two may be related. Like CVS, migraine is characterised by headaches that begin and end abruptly, followed by longer periods without pain or other symptoms. They both have a number of similar triggers, including stress and excitement. Many children with CVS have a family history of migraine or suffer migraine attacks as they grow older. Given these similarities, CVS patients subject to frequent and long-lasting episodes are treated with some success using propranolol, cyproheptadine and amitriptyline, which are generally used for treating migraine headaches.

and propel it towards the colon. The small intestine is 6–7 metres long in an adult. It is divided into three sections, namely the:

1. duodenum, the first 0.25 metres of the small intestine
2. jejunum, the next 1.5–2 metres
3. ileum, the remaining 3.5 metres

Digestion occurs primarily in the duodenum and jejunum and absorption mainly in the ileum. The duodenal walls contain goblet cells that secrete mucus that protects it against damage by acidic chyme. Crypts of Lieberkühn are pits within the wall of the small intestine with secretory cells that release intestinal juice and Paneth cells that secrete the antibacterial enzyme lysozyme. Intestinal juice is largely water with mucus and buffered to a pH of about 7.6. Along with pancreatic juice (see below), it neutralises chyme and provides a liquid medium that aids the absorption of nutrients. The epithelial lining consists of enterocytes that cover 0.5–1.5 mm-long cylindrical projections into the lumen called villi (*Figure 11.17*). In turn, each enterocyte has its apical plasma membrane surface, that is, the side facing the lumen of the GIT, arranged into microscopic extensions 1 µm long and 0.1 µm in diameter called microvilli that form a **brush border** coating the villi (*Figure 11.18*). The opposing face of the enterocyte is the basolateral membrane. Villi and microvilli greatly increase the internal surface area of the intestine. Villi are self-renewing structures. Undifferentiated stem cells (*Chapters 6* and *17*) located in the crypts divide constantly to provide a constant supply of new cells that migrate along a villus to its tip where they are shed into the GIT lumen. During this migration, the cells differentiate and form mature enterocytes. This constant shedding of cells and their replacement means the entire epithelial lining is renewed approximately every six days. Hydrolytic enzymes, for example the enzymes that hydrolyse disaccharides (*Margin Note 11.3*), are integral proteins of the plasma membranes of the microvilli.

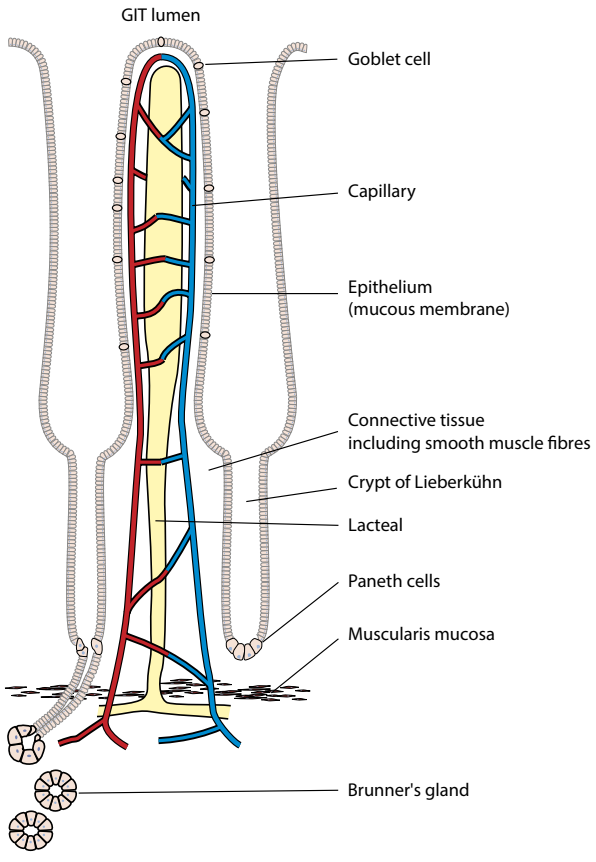


Figure 11.17 A gastrointestinal tract (GIT) villus with its associated crypts of Lieberkühn.

Margin Note 11.3
Disaccharides

Disaccharides, as their name implies, are composed of two simple sugars (monosaccharide residues) covalently bound together by a glycosidic bond. A variety of disaccharides are known, which can be found naturally or produced by the action of digestive enzymes on polysaccharides. Disaccharides differ from one another in their constituent monosaccharides and the type of glycosidic bond joining them together (*Table 11.1*).

Margin Note 11.4
Zymogens

Digestive enzymes are produced as inactive precursors called zymogens (*Table 11.2*), which ensures the protection of the cells and tissues that produce them from the catalytic activities of the enzymes. Activation is achieved by a partial and specific hydrolysis of part of the zymogen structure that unmasks the active site.

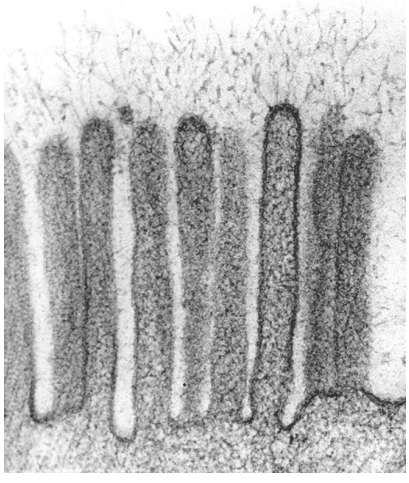


Figure 11.18 Electron micrograph of microvilli of the gastrointestinal tract. Courtesy of Emeritus Professor D.W. Fawcett.

Disaccharide	Structure	Digestive (hydrolytic) enzyme	Products of digestion (hydrolysis)
Isomaltose (product of starch digestion)	Glucose(α,1→6) glucose	Isomaltase	2 glucoses
Lactose (milk sugar)	Galactose(β,1→4) glucose	Lactase	Galactose + glucose
Maltose (product of starch digestion)	Glucose(α,1→4) glucose	Maltase	2 glucoses
Sucrose (common sugar of plants)	Glucose(α,1→β,2) fructose	Sucrase	Fructose + glucose
Trehalose (found in fungi and some insects)	Glucose(α,1→1) glucose	Trehalase	2 glucoses

Table 11.1 Some common disaccharides, their structures (glycosidic bonds are indicated in parentheses), associated digestive enzymes and products of their digestion. ⏎

Pancreatic juice contains water, alkaline salts that give it a pH of 7.8 to 8.0, enzymes and zymogens (*Margin Note 11.4*), and enters the small intestine through the pancreatic duct. The salts neutralise acid from the stomach (*Box 11.5*).

Proteolytic enzymes can be divided into exopeptidases and endopeptidases (*Figure 11.21*). Exopeptidases are aminopeptidases situated in the apical membranes of enterocytes and carboxypeptidases that are components of pancreatic juice. Both types of enzymes catalytically remove amino acids from the respective ends of proteins and

BOX 11.5 ACID AND ALKALI ⏎

Parietal cells are found mainly in the gastric fundus. These cells have microvilli-lined tubular structures derived from the endoplasmic reticulum called canaliculi, which form a complex system that greatly increases the surface area available for acid secretion. *Figure 11.19* illustrates the production of HCl. The dissociation of water generates H^+ that are actively transported out of the luminal surface of the cell by an H^+/K^+-ATPase, commonly referred to as a H^+ (proton) pump, using free energy from the hydrolysis of ATP to drive the process. Parietal cells are rich in carbonic anhydrase that catalyses the formation of H_2CO_3, which can react with the OH^- from the dissociation of water to regenerate water and HCO_3^-. The latter is exchanged for Cl^- by a HCO_3^-/Cl^- antiporter in the basolateral membrane. The Cl^- can leave the cell at the luminal surface via a K^+/Cl^- symport. Thus, K^+ shuttles in and out of the cell in exchange for H^+ and Cl^-, respectively, although Cl^- can also leave through a uniport. The secretion of H^+ is an energy-intensive process and parietal cells contain a substantial number of mitochondria to furnish the necessary ATP. The initial HCl secreted has a pH of approximately 0.8 and is delivered into the gastric lumen through the mucus layer where it is diluted to a working pH of 1–2.

In contrast with parietal cells, duct cells of the exocrine pancreas produce an alkaline secretion of pH 7.8–8.0 (*Figure 11.20*). Carbonic anhydrase catalyses the formation of H_2CO_3, which dissociates to form HCO_3^- and H^+. Hydrogen carbonate also enters the cell from the interstitial fluid via a Na^+/HCO_3^- symport in the basal plasma membrane. The exchange of HCO_3^- for Cl^- at the luminal surface and the paracellular secretion of Na^+ between the cells renders the duct contents alkaline. Any excess Na^+ is pumped out of the cell by the Na^+/K^+-ATPase (*Section 11.6* and see *Figure 8.10*).

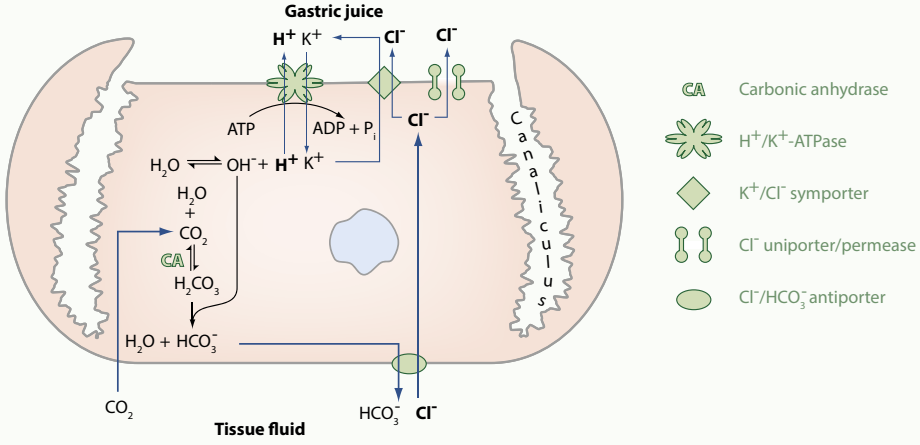


Figure 11.19 Overview of the formation of HCl by a gastric parietal cell. ⏎

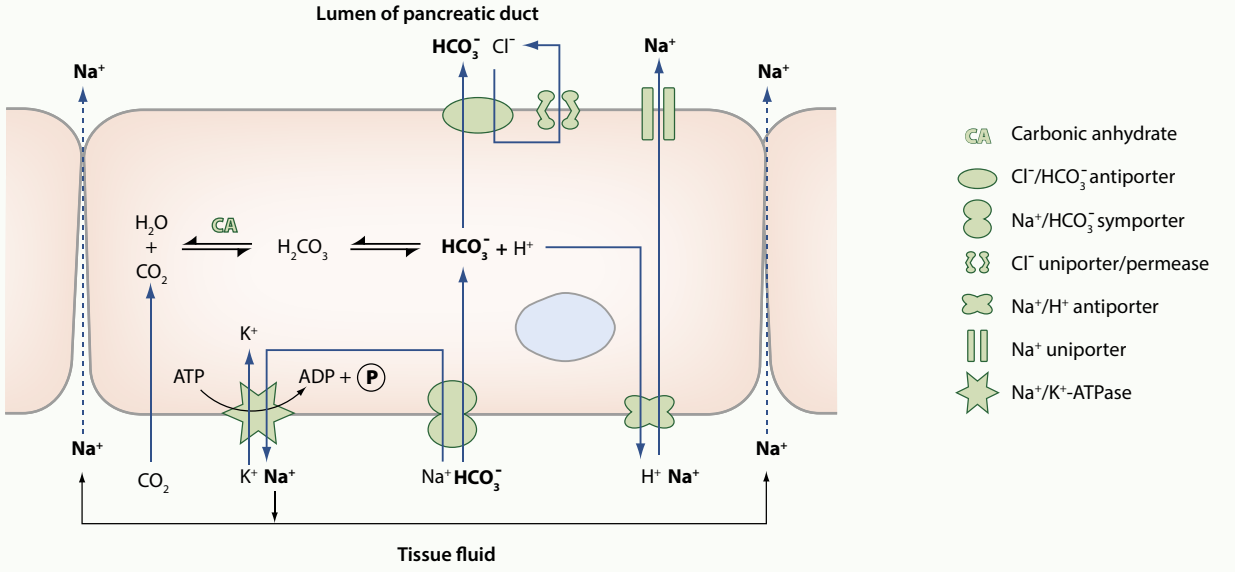


Figure 11.20 Simplified illustration of the secretion of $NaHCO_3$ by a pancreatic duct cell.

Zymogen	Activation mechanism	Active enzyme
Pepsinogen	HCl, pepsin	Pepsin
Trypsinogen	Enteropeptidase (enterokinase)	Trypsin
Proelastase	Trypsin	Elastase
Chymotrypsinogen	Trypsin	Chymotrypsin
Procarboxypeptidase	Trypsin	Carboxypeptidase
Proaminopeptidase	Trypsin	Aminopeptidase

Table 11.2 Zymogens.

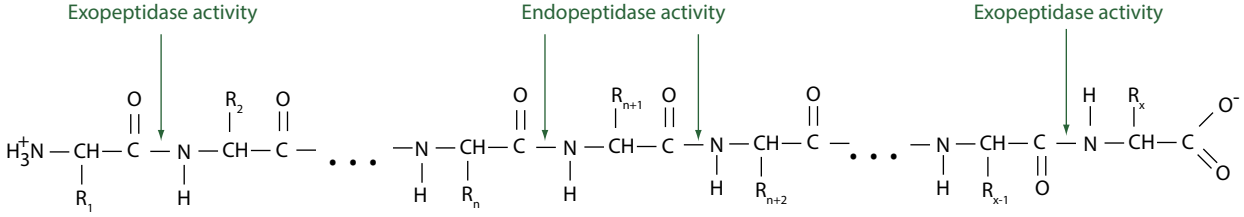


Figure 11.21 The sites of digestion (hydrolysis) of proteins by exopeptidase and endopeptidase activities.

peptides. The pancreatic proteases chymotrypsin, elastase and trypsin are endoproteases and catalyse the hydrolysis of proteins and peptides at different peptide bonds forming smaller peptides and amino acids. The actions of endopeptidases increase the number of protein ends, effectively raising the concentration of substrates for exopeptidase action. Pancreatic amylase continues the digestion of polysaccharides to produce disaccharides.

Pancreatic lipase (*Figure 11.22*) hydrolyses lipids to fatty acids and 2-monoacylglycerols but requires the presence of colipase, another protein secreted by the pancreas, for full activity. Bile released from the gallbladder enters the intestine through the bile duct and emulsifies lipids into droplets of about 1 μm diameter, which greatly increases the surface area available for lipase activity.

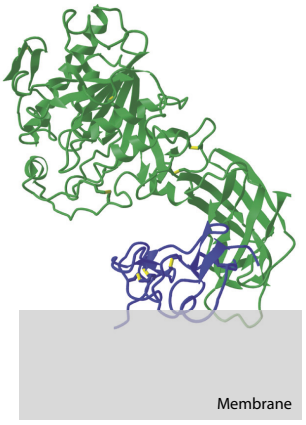


Figure 11.22 Molecular model of a pancreatic lipase (green) and colipase (blue) complex. PDB file 1N8S.

Pancreatic esterase activity can complete the digestion to glycerol and free fatty acids (*Figure 11.23*).

The enzymes concerned with the final stages of digestion of some foods are located in the brush border of the enterocytes or even within their cytoplasm. This ensures that the final products of digestion are produced near or within the absorptive surface of the GIT.

11.6 ABSORPTION OF THE PRODUCTS OF DIGESTION

The surface area of the interior of the small intestine is estimated to be 300 m^2 in humans. This substantial area is approximately 600 times greater than a uniform tube of comparable diameter and is achieved by a number of adaptions. Numerous deep circular folds, also called valves of Kerckring, triple the available surface area while mixing the contents and ensuring it travels in a spiral manner that slows the rate of flow and gives more time for digestion and nutrient absorption. However, most of the increase in area is attributable to the villi and microvilli described in *Section 11.5*, which increase the absorptive area by 10- and 20-fold, respectively. These structural adaptions facilitate a rapid absorption of nutrients. Chyme is mixed and moved along the duodenum by peristaltic movements that begin in the duodenum and travel up to approximately 700 mm before ceasing. However, before this cessation occurs, another peristaltic wave begins starting at a position in advance of the first. Thus, chyme is propelled forwards by successive, overlapping waves during which digestion and nutrient absorption are completed.

Enterocytes are joined together by tight junctions that ensure that material, in general, cannot leak from the lumen. Thus, the absorption of nutrients by enterocytes is largely active and selective and requires considerable amounts of metabolic (free) energy. A major proportion of this energy is obtained from a membrane-bound Na^+/K^+-ATPase (see *Figure 8.10*) that catalyses the hydrolysis of ATP when expelling three Na^+ from the cell and pumping two K^+ into it:

$$3Na^+_{IN} + 2K^+_{OUT} + ATP + H_2O \rightarrow 3Na^+_{OUT} + 2K^+_{IN} + ADP + P_i$$

Since both Na^+ and K^+ are being transported against their electrochemical gradients, both movements are active transport. Also, since more positive charges are being pumped out of the cell than are entering it, the effect contributes to the potential difference across the membrane, called the **resting membrane potential**, of about –60 mV: the inside of a cell being negative with respect to the outside. The membrane potential of the apical membrane of the enterocyte can be harnessed to facilitate the uptake of a variety of nutrients from the lumen of the GIT. Both monosaccharides and amino acids are transported across the enterocyte luminal membrane in a Na^+-dependent fashion. A variety of different membrane transporter proteins are responsible for the absorption of nutrients.

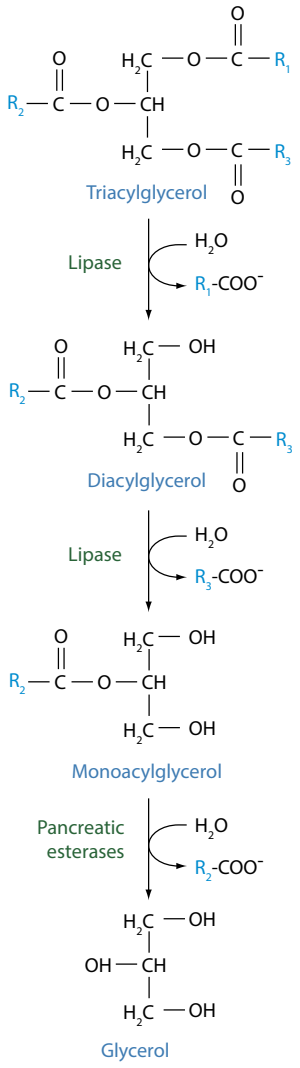


Figure 11.23 Overview of the digestion (hydrolysis) of triacylglycerols.

Carbohydrates

Glucose and galactose are transported across the enterocyte luminal membrane in an active, Na^+-dependent fashion by the same transporter. One molecule of these sugars can only move through the transporter into the cell if a Na^+ moves in at the same time (*Figure 11.24*). The concentrations of the sugars can build up within the cytoplasm, such that they are able to leave the cell through the basolateral membrane by facilitated diffusion. Other monosaccharides, for example fructose and trehalose, are absorbed only by *facilitated* diffusion and are absorbed to a much lesser extent.

Peptides and amino acids

The major initial products of protein digestion are small peptides. Tetrapeptides are hardly taken up, with most absorption concentrated on di- and tripeptides. These are absorbed over the luminal surfaces of enterocytes of the jejunum by the peptide transporter protein PepT1 in an H^+-dependent fashion (*Figure 11.25*) that mimics the uptake of glucose and galactose. Within the cytoplasm, the peptides are hydrolysed to amino acids, ensuring a continuous sink is present to facilitate peptide uptake by the cells. The exit of the amino acids from the cells on the basolateral side also occurs down their concentration gradients. However, as peptides are moved further along the GIT, they are hydrolysed by peptidases to free amino acids and their absorption occurs in the ileum using a number of Na^+-dependent transporters (*Figure 11.25*), which have specificities for different amino acid side chains. Peptides can also be absorbed by a paracellular route where they pass between enterocytes, rather than being absorbed across the luminal surface. Relatively large peptides can be absorbed by this method and may initiate immune reactions leading to food allergies (*Chapters 5* and *10* and *Box 11.6*).

Margin Note 11.5
Pancreatic lipase and cystic fibrosis

Patients suffering cystic fibrosis (CF; *Chapter 16*) are defective in pancreatic lipase and other enzyme activities due to exocrine pancreatic insufficiency (*Section 11.11*), when the pancreatic duct becomes blocked with a thick secretion. This reduces their ability to digest lipids and leads to lipid malabsorption and steatorrhoea (*Section 11.10*). Many CF patients are treated with supplements of porcine pancreatic extracts that contain pancreatic lipase activity. However, these supplements are needed in high doses because the duodenums of many CF patients are more acidic than in healthy individuals and denaturation or inhibition of the enzyme can occur. The treatment is effective in only 30% of CF patients and 15–20% still suffer from steatorrhoea.

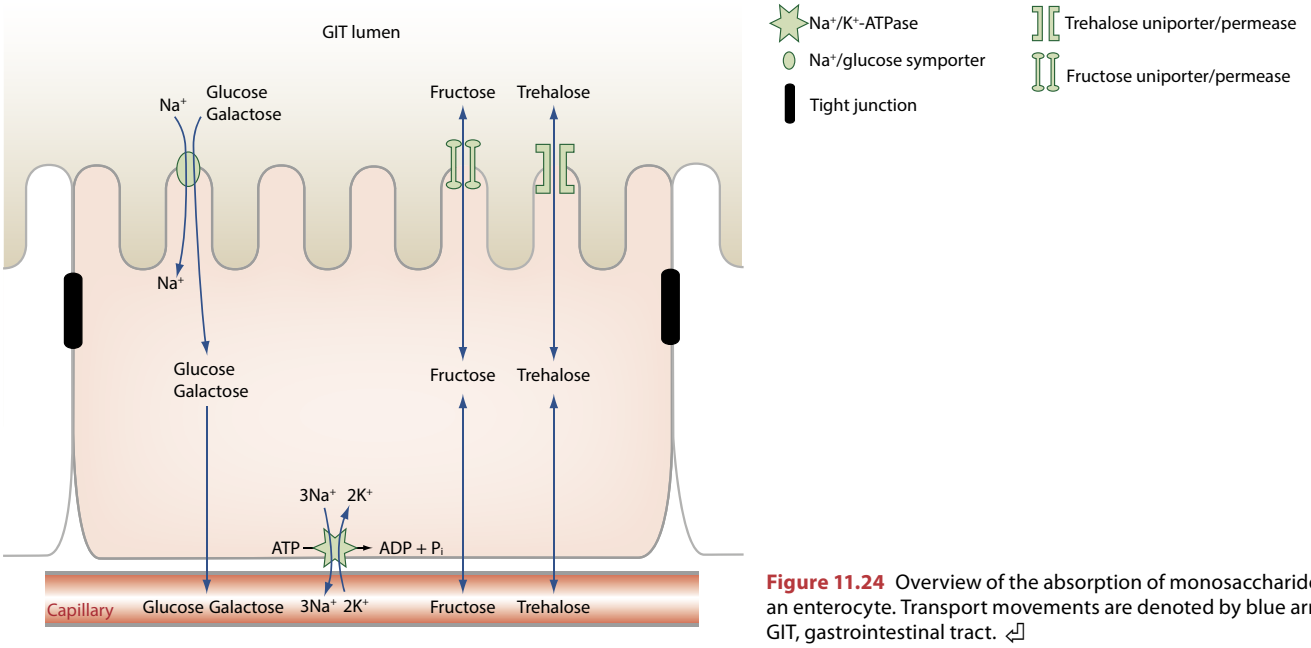


Figure 11.24 Overview of the absorption of monosaccharides by an enterocyte. Transport movements are denoted by blue arrows. GIT, gastrointestinal tract.

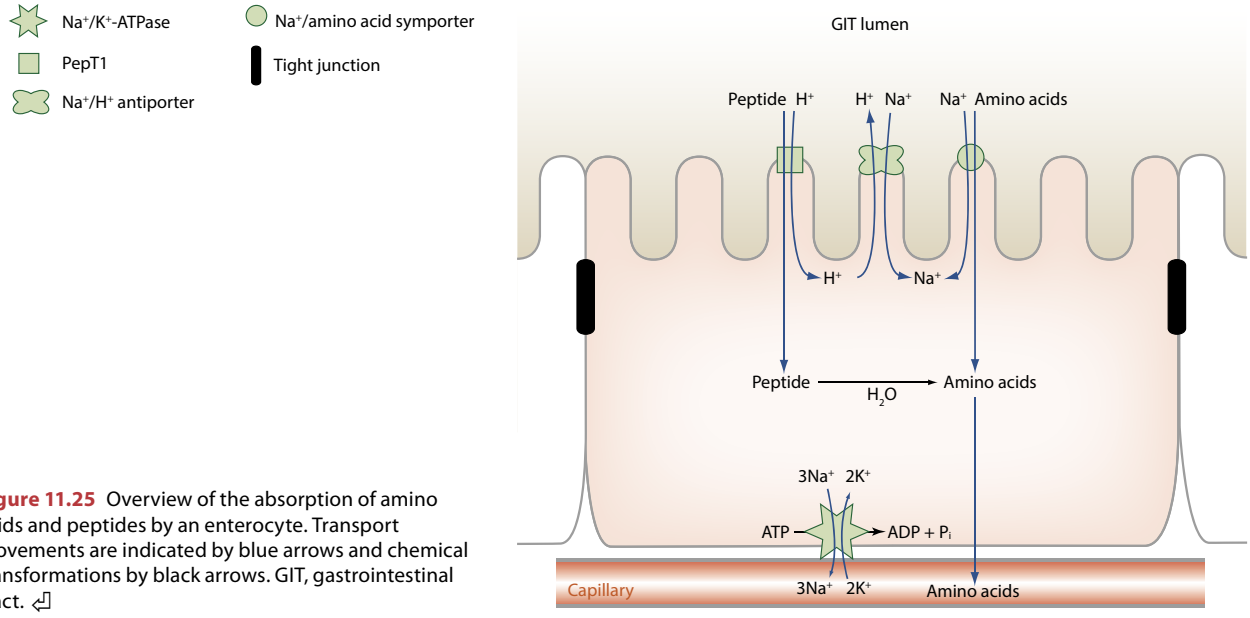


Figure 11.25 Overview of the absorption of amino acids and peptides by an enterocyte. Transport movements are indicated by blue arrows and chemical transformations by black arrows. GIT, gastrointestinal tract.

Nucleosides and nucleotides

Digestion of RNA produces nucleotides that are further degraded to nucleosides at the brush border, which, again, are absorbed in a Na^+-dependent manner. Catabolism within the cytoplasm converts the nucleotides to ribose phosphate and bases. Eventually, the purine bases are converted to urate and the pyrimidines to uracil as shown in *Figure 11.26*.

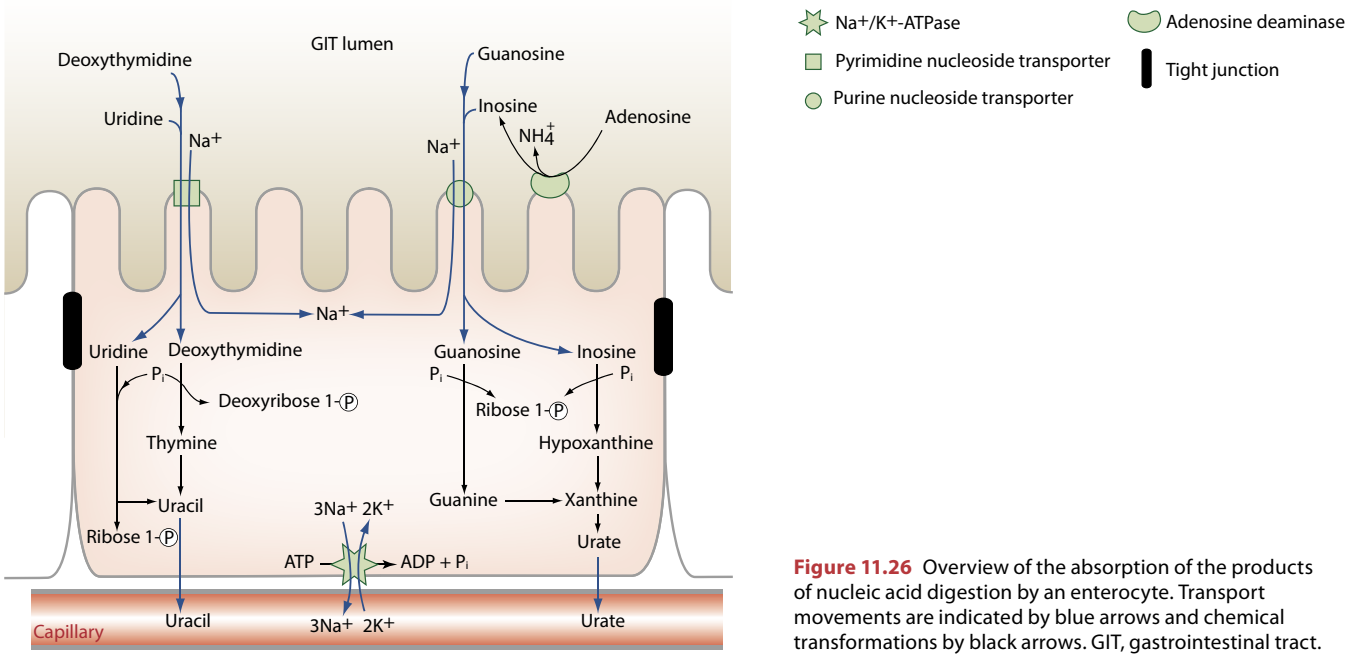


Figure 11.26 Overview of the absorption of the products of nucleic acid digestion by an enterocyte. Transport movements are indicated by blue arrows and chemical transformations by black arrows. GIT, gastrointestinal tract.

Lipids

Fatty acids, monoacylglycerols, monoacylphospholipids and cholesterol are absorbed as mixed micelles of approximately 6 nm diameter by the brush border of the enterocytes (*Figure 11.27*). Triacylglycerols and phospholipids are then reformed within the enterocyte cytoplasm and packaged into lipoprotein complexes called chylomicrons of 75–1000 nm diameter (*Chapter 14*), which are transported from the GIT in lacteals of the lymphatic system. This ensures that the lipids bypass the liver and are delivered to the blood through the thoracic duct.

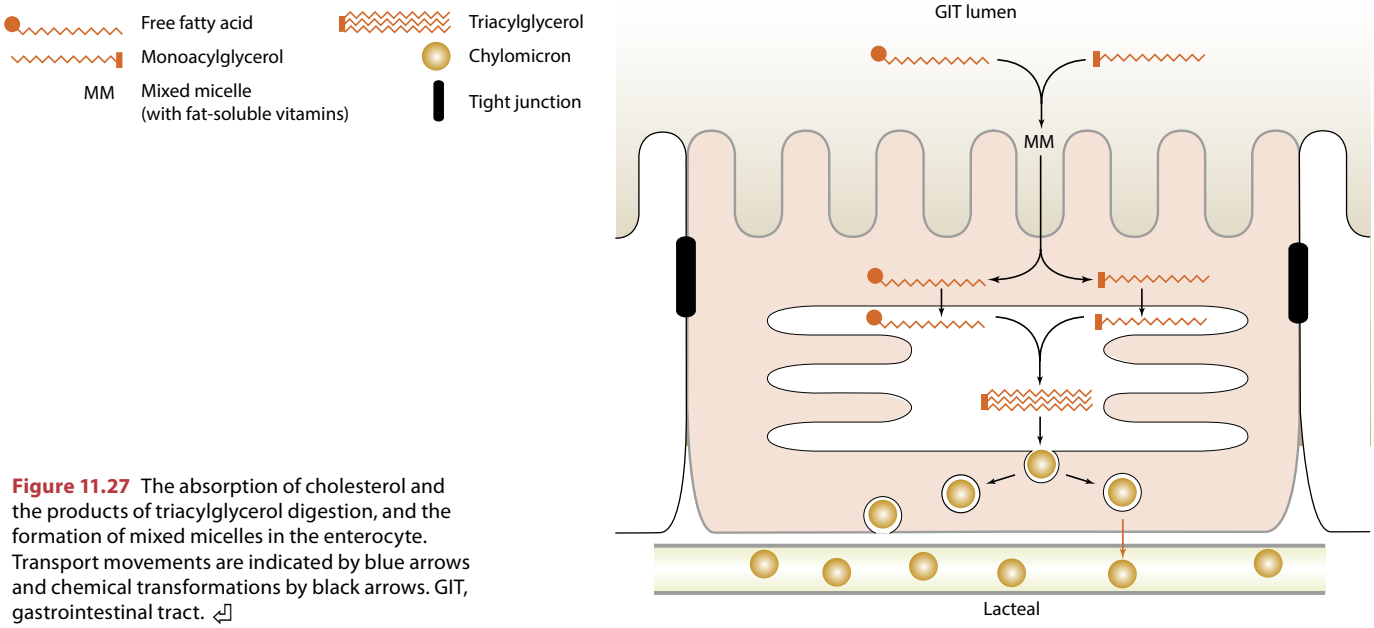


Figure 11.27 The absorption of cholesterol and the products of triacylglycerol digestion, and the formation of mixed micelles in the enterocyte. Transport movements are indicated by blue arrows and chemical transformations by black arrows. GIT, gastrointestinal tract.

Vitamins

Water-soluble vitamins are generally actively absorbed by enterocytes using specific transporter proteins. Vitamin B_1 (thiamine) occurs in foods normally as phosphate derivatives that are hydrolysed by intestinal phosphatases releasing the free vitamin. This is then absorbed in a Na^+-dependent fashion like that of glucose (see *Figure 11.24*) by thiamine transporter proteins (THTRs) 1 and 2 found throughout the small intestine. Given only eggs and milk contain appreciable quantities of free vitamin B_2 (riboflavin), the vitamin is mainly freed for absorption by the digestion of flavoproteins and the subsequent hydrolysis of the released coenzymes by phosphatases. Uptake occurs in a Na^+-dependent manner in the upper portions of the small intestine via riboflavin transporters (RFVTs) 1, 2 and 3. Vitamin B_2 is also biosynthesised in the colon where it is absorbed in a carrier-dependent fashion. Vitamin B_3 (niacin, nicotinic acid and nicotinamide) is found in foods mainly in the form of coenzymes in proteins and is released when they are digested. The free vitamers are then absorbed in an H^+-dependent manner that resembles the uptake of dipeptides (see *Figure 11.25*), rather than being Na^+-dependent.

The vitamins B_5 (pantothenate) and B_7 (biotin or vitamin H) share only a superficial structural resemblance. However, both are actively absorbed into enterocytes by the Na^+-dependent multivitamin transporter (SMVT), which is also responsible for the absorption of the antioxidant lipoic acid (*Figure 11.28*).

Vitamin B_6 (pyridoxine, pyridoxal, pyridoxamine) phosphates supplied in the diet are hydrolysed by phosphatases of the intestinal mucosa with the release of the free vitamers. These appear to be absorbed by passive diffusion, although some studies indicate that, like vitamins B_5 and B_7, they are also absorbed by an H^+-dependent carrier. Vitamin B_6 is also biosynthesised by bacteria in the colon and absorbed by colonocytes.

In the diet, vitamin B_9 (folate or vitamin M) occurs conjugated with polyglutamates. Folate conjugase activity releases the vitamin from the complex in the upper small intestine. This allows its absorption by one of two carriers: the reduced folate carrier (FOLT) and an H^+-coupled folate transporter (PCFT), which transfer folate across the enterocyte membrane. Within the cell, the vitamin is reduced and methylated. Hence, it is largely methyltetrahydrofolate that enters the bloodstream. Folate synthesised by bacteria of the large intestine is also a source of the vitamin and is absorbed by colonocytes.

In the stomach, vitamin B_{12} (cobalamin) binds to haptocorrin (see *Figure 11.16*). It is released in the small intestine by the action of trypsin, where it forms a complex with GIF (*Figure 11.29*). Uptake occurs through the activities of the anion-associated transmembrane protein and cubilin (the intestinal receptor for GIF) of the ileal mucosa. This complex, sometimes referred to as 'cubam', is the receptor for the cobalamin (B_{12})–GIF complex, which enters the enterocyte by endocytosis and is processed through the endosomal lysosome system (*Chapter 16*). The released free cobalamin enters the circulation through the basolateral membrane and complexes with a further binding protein, transcobalamin (see *Figure 13.31*), and is transported to the liver.

Vitamin C (ascorbate) absorption requires the activity of two Na^+-coupled vitamin C transporters (SVCT1 and -2). Its oxidised form, dehydroascorbate, is transferred from enterocytes through glucose transporters in the basolateral membrane.

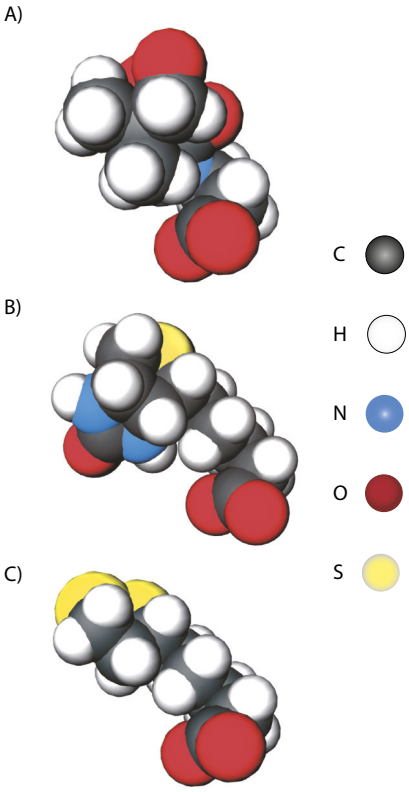


Figure 11.28 Computer drawn models of vitamins (A) B_5 (pantothenate), (B) B_7 (biotin or vitamin H) and (C) lipoate.

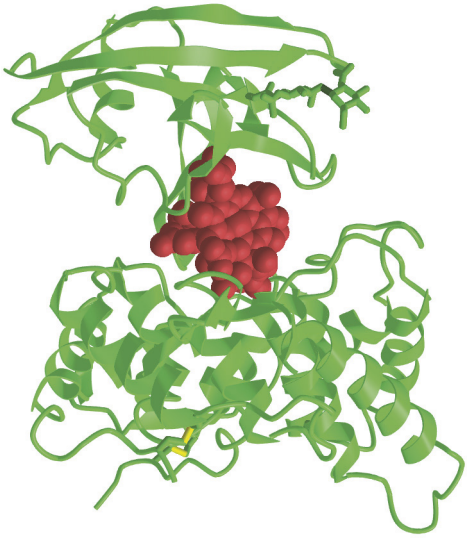


Figure 11.29 Molecular model of gastric intrinsic factor in green, with its bound vitamin B_{12} (cobalamin) highlighted in red. PDB file 2PMV.

The fat-soluble vitamins, A, D, E and K, are generally thought to be absorbed by passive diffusion within the mixed micelles of the fatty acids monoacylglycerols, monoacylphospholipids and cholesterol described above (see *Figure 11.27*). For these reasons, a deficiency in dietary lipids means that the absorption of fat-soluble vitamins is greatly reduced. However, the apical membrane transporter SR-B1, encoded by *SCARB1*, provides an alternative route for diffusion across the membrane. This integral membrane protein forms a hydrophobic lined pore that traverses the membrane, forming a tunnel for the movements of a variety of lipids, including lipid vitamins and carotenoids. Lipid-soluble vitamins exit the enterocyte across the basolateral membrane in chylomicrons.

Minerals and trace elements

Many minerals and trace elements are absorbed into enterocytes using specific transport proteins. Here, they bind to cytoplasmic proteins before being transferred across the basolateral membrane in an energy-dependent manner. Transfer into the blood, however, only occurs when there are sufficient free plasma binding proteins to accept the ions. Uptake of some ions can also occur by paracellular routes.

Most Ca^{2+} is absorbed in the upper reaches of the small intestine. When abundant in the diet, the majority is absorbed by diffusion through the tight junction, that is, paracellularly, although the protein composition of the junction can modify the rate of uptake. At lower concentrations, Ca^{2+} enters the enterocyte through the Ca^{2+} channel, TRPV6 (transient receptor potential cation channel) and, to a lesser extent, TRPV5, moving in accordance with its electrochemical gradient. Within the cytoplasm it binds to calbindin before leaving the cell actively at the basolateral membrane through a Ca^{2+}-ATPase (Ca^{2+} pump) and to a lesser extent via an antiporter in exchange for Na^+ (*Figure 11.30*). The 1,25 $(OH)_2D_3$ form of vitamin D increases epithelial permeability and stimulates absorption and extrusion. Approximately 50% of Ca^{2+} in the blood is the free ion; the remainder is bound to serum albumin or chelated to citrate.

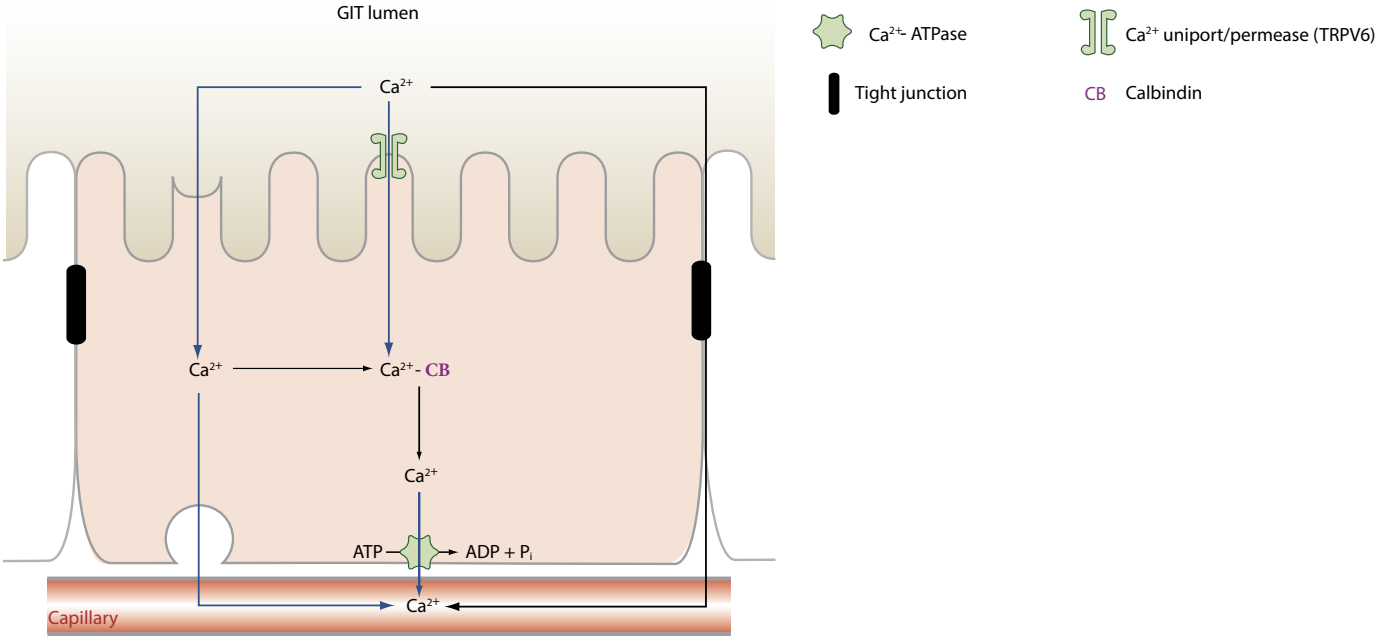


Figure 11.30 Schematic outline of Ca^{2+} absorption by an enterocyte. GIT, gastrointestinal tract.

Copper absorption occurs primarily in the duodenum. It is first reduced from its cupric (Cu^{2+}) to the cuprous (Cu^+) state by the activity of an apical membrane-bound cytochrome *b* reductase. The Cu^+ is then transferred into the enterocyte by an apical

membrane copper transporter 1 (CTR1). Copper toxicity in the cytosol is avoided by its transfer to Cu transporting ATPases in the Golgi apparatus membrane or in the basolateral membrane via a series of binding proteins. Mutations that impair the activity of the Golgi apparatus ATPase result in Menkes disease (*Chapter 10*). Transfer across the basolateral membrane only occurs if available ceruloplasmin is present in the plasma to transport the copper (*Figure 11.31*).

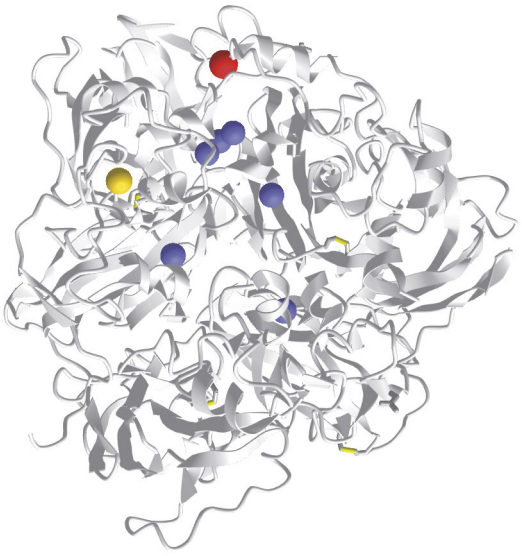


Figure 11.31 Molecular model of ceruloplasmin. Six bound Cu^{2+} are shown as blue spheres. Bound Ca^{2+} and Na^{+} are shown as red and yellow spheres, respectively. PDB file 4ENZ.

Dietary iron occurs in nonhaem and haem forms and their absorption by duodenal enterocytes is by different mechanisms. Dietary nonhaem iron occurs in both ferrous (Fe^{2+}) and ferric (Fe^{3+}) forms. Ferrous iron is soluble up to a pH of about 7, but the predominant form, Fe^{3+}, is sparingly soluble above pH 3 and is not available for absorption and must be reduced before it can be transported across the intestinal epithelium. Ascorbate in the diet and a cytochrome *b* reductase (Dcytb) reduces Fe^{3+} to Fe^{2+}. Ferrous iron is then transported into the enterocyte through the divalent metal transporter 1 (DMT1) in an H^{+}-coupled mechanism (*Figure 11.32*). This transporter can also affect the absorption of other divalent metal ions, such as those of cadmium, cobalt copper, lead, nickel and zinc. Haem iron is absorbed by a haem receptor and, once internalised, its ferrous iron is released into the intracellular pool by haem oxygenase activity (*Figure 11.32*). Ferrous iron is exported from the enterocyte across the basolateral membrane by the membrane protein ferroportin l or Ireg1. It is then oxidised by hephaestin, a transmembrane copper dependent ferroxidase, which is necessary for effective iron transport. The Fe^{2+} is bound by transferrin (*Figure 11.33*) in the plasma for transport and can be stored by erythrocytes in ferritin molecules (*Chapter 13*).

The uptake of Na^{+} has already been mentioned in relation to the active uptake of several nutrients and many anions, hydrogen carbonate, chloride and iodide, can passively follow it into enterocytes. Phosphate is actively accumulated by enterocytes in the upper portions of the small intestine in a Na^{+}-dependent manner using a Na/P_i antiporter. However, at high concentrations, it can be absorbed through the paracellular pathway.

Water

Approximately 9 L of fluid pass through the GIT each day. Reabsorption of water from the GIT is essential to prevent dehydration. Most, about 95%, is absorbed by the small intestine, while 4% is retained by the large intestine and only 1% is lost from the GIT (see *Box 11.8*).

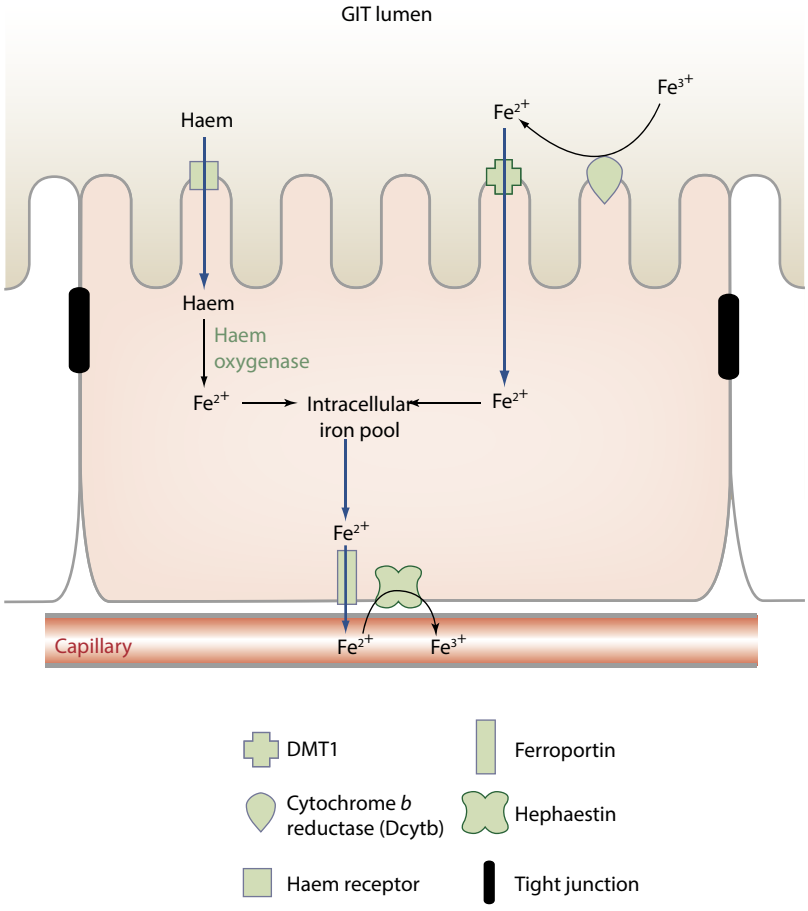


Figure 11.32 Uptake of iron by an enterocyte. Transport movements are denoted by blue arrows and chemical transformations by black arrows. GIT, gastrointestinal tract.

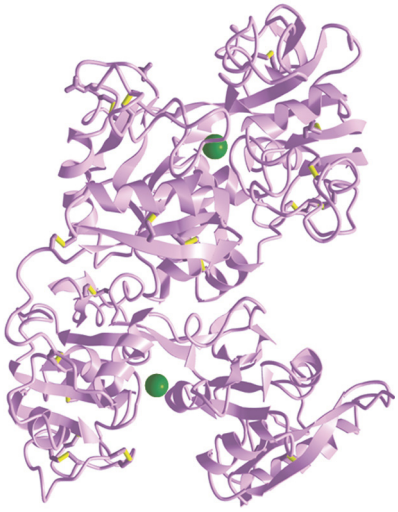


Figure 11.33 Molecular model of transferrin, with its two bound Fe^{2+} shown as green spheres. PDB file 3QYT.

BOX 11.6 FOOD ALLERGIES AND INTOLERANCE

Food allergies are immune reactions to food components that are allergenic to the individual in question (*Chapter 5*). Foods commonly associated with allergies include milk, eggs, peanuts, tree nuts, soy, wheat, fish and shellfish, which are protein-rich and account for over 90% of all food allergies. Food allergies are relatively common and affect 1–2% of adults and up to 8% of children in the UK. The antibodies responsible for food allergies are immunoglobulin E (IgE) molecules that react with allergens to trigger the release of histamine (*Figure 11.34*). Histamine is largely responsible for the symptoms, but these differ between patients and can differ in the same individual when exposed to different allergens. Symptoms include skin irritations, for example rashes and eczema, itchy nose and eyes, sneezing, excessive nasal mucus, coughing and shortness of breath, nausea, vomiting and diarrhoea. Skin rashes, such as nettle rash (also called urticaria or hives), can appear but tend to last only a few days. Occasionally, patients present with long-lasting, chronic skin reactions such as scaly patches. Some individuals may experience a severe anaphylactic reaction as described in *Chapter 5*. This is triggered by a small amount of food and is a rare but potentially fatal condition in which several different parts of the body experience allergic reactions including skin rashes, swollen throat and difficulties in breathing, nausea, vomiting, diarrhoea, hypotension and unconsciousness. Symptoms usually appear within minutes of exposure to the allergen and can last several hours. An unusual form of anaphylaxis occurs following the consumption of allergens within a few hours of exercising and is referred to as exercise-induced anaphylaxis.

Diagnosing a food allergy involves taking a clinical history from the patient, focusing on past allergic reactions and considering seasonal or environmental cues. This is followed by a clinical examination to detect characteristic signs and symptoms affecting the eyes, skin and nose. Skin prick testing is often performed. The procedure introduces a small amount of allergen into the skin and is generally safe. A tiny puncture, usually in the forearm, is made with a lancet and a drop of allergen in solution or suspension is placed on the opening. A positive reaction to the allergen is indicated by itching at the site within a few minutes, followed by reddening, swelling and a raised weal at the centre. The weal enlarges to a maximum size within 15–20 minutes, when it is measured and the size recorded. A negative response indicates that the patient is insensitive to the allergen under test.

Blood tests, such as the radioallergosorbent test (RAST), to detect the presence of IgE antibodies may be used, but can only indicate an allergy if the specific IgE is present. However, there is no clear relationship between the concentration of blood IgE and the severity of the allergy. Occasionally, a challenge test may be performed, where the suspected food is given to the patient first by touch and then by ingestion in

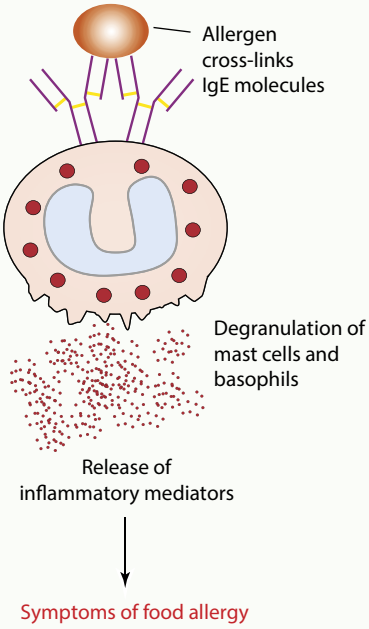


Figure 11.34 Schematic showing the release of inflammatory mediators in response to a food allergy. See *Chapter 5* for further details. IgE, immunoglobulin E.

increasing amounts and the response monitored. Challenge tests should always be performed under supervised conditions in a hospital or clinical setting, so that any clinically serious reactions can be detected and managed.

Food allergies are usually treated by avoiding the offending food. In some cases, complete elimination of the food for one to two years results in a loss of clinical symptoms, although certain allergies, for example those associated with the intake of fish or peanuts, may persist for life. It is vital that consumers are provided with clear information on the composition of foods they purchase. Food avoidance may have serious nutritional consequences, particularly when a key nutrient is removed from the diet. Immediate medical attention is necessary when anaphylaxis occurs and usually involves an injection of adrenaline to dilate airways and blood vessels by relaxing smooth muscle tissues.

Food intolerance refers to the less well-defined condition occurring in some people when they eat certain foods and is characterised by headaches, muscle and joint aches and pains, and tiredness. The less well-defined symptoms make the condition harder to diagnose than classical food allergy. Examples of food intolerances include coeliac disease and lactose intolerance.

11.7 SMALL INTESTINE AND HOMEOSTASIS

Within enterocytes, a portion of the monosaccharides absorbed are converted to lactate by glycolysis. Excess nonessential amino acids, especially glutamine, are used to synthesise alanine and ammonia (*Figure 11.35*). These products are then delivered to the liver in the hepatic portal vein. Converting some of the absorbed nutrients to lactate and alanine

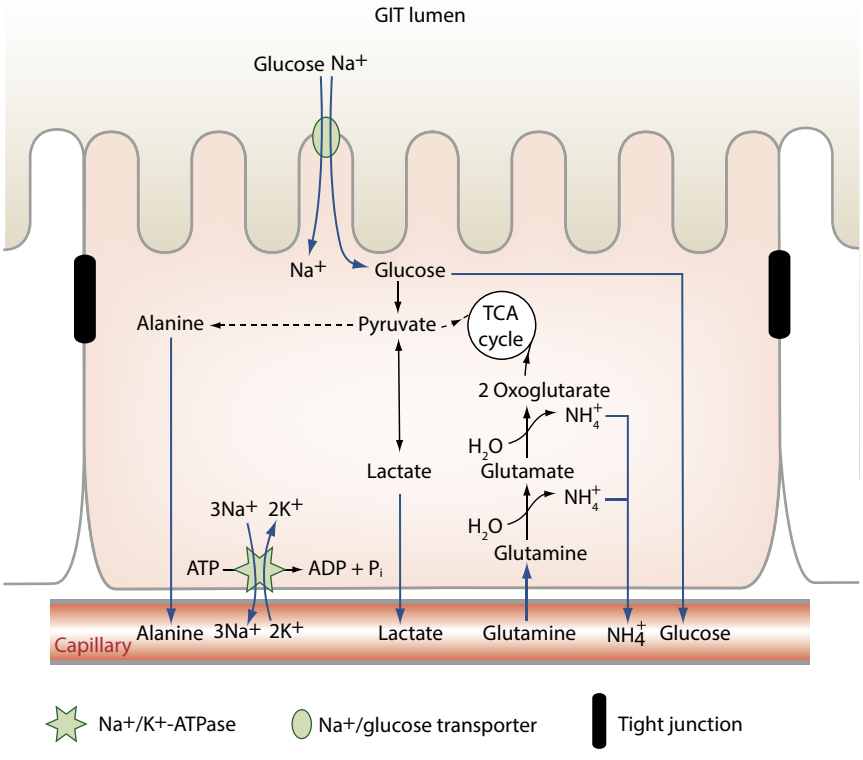


Figure 11.35 Overview of the role of the enterocyte in energy homeostasis. Transport movements are indicated by blue arrows and chemical transformations by black arrows. GIT, gastrointestinal tract.

reduces the metabolic load on the liver because it can easily regenerate pyruvate from them. Pyruvate is a versatile liver metabolite; it is a substrate for the TCA cycle, allowing the formation of ATP during oxidative phosphorylation but it can be used for the biosynthesis of glucose and glycogen, ketone bodies, fatty acids, all but two of the nonessential fatty acids and also cholesterol. The GIT is a significant contributor to nutrient homeostasis both during and after nutrient absorption because the formation of lactate and alanine continues even when absorption ceases.

11.8 LARGE INTESTINE

The large intestine is so named because its diameter is greater than that of the small intestine, although it is, in fact, much the shorter of the two. Fluid, containing the unabsorbed products of digestion, directly enters from the small intestine at a junction that is also the site of the vestigial caecum and appendix (*Figure 11.36*). Absorption of Na^+ and water occurs over the surface of the large intestine, which also acts as a reservoir for material resistant to digestion by GIT enzymes. However, colonic bacterial action on this material synthesises some nutrients, for example a number of vitamins (*Section 11.6*) and about 200–2000 mL of gas (flatus), which is released in 10–14 episodes per day. The final waste together with bacteria forms the faeces, which pass to the last section of the GIT, the rectum, and is eliminated through the anus. Two sphincter muscles control elimination: the first of smooth muscle opens involuntarily in response to pressure within the rectum; the second is controlled voluntarily and allows for a conscious decision to defaecate.

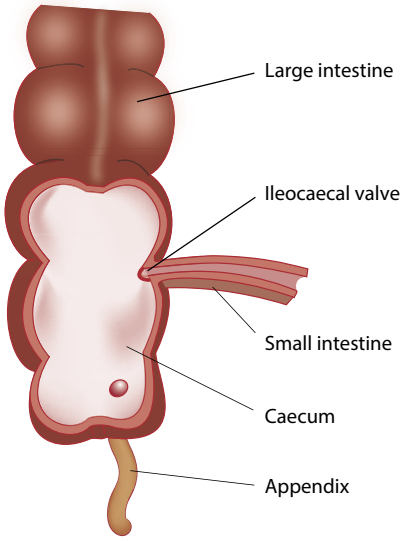


Figure 11.36 The junction of the small and large intestines.

11.9 HORMONAL CONTROL OF GASTROINTESTINAL TRACT SECRETIONS

The GIT produces a variety of hormones, many of whose functions are not well understood, although some of them, together with neuronal activities, are concerned with coordinating the secretions of various digestive juices. Endocrine cells are scattered throughout the entire GIT in clusters forming a diffuse portion of the endocrine system (*Chapter 7*).

The G cells in the antral and pyloric regions of the stomach produce gastrin. Gastrin occurs in a number of molecular forms; for example, $gastrin_{17}$ and $gasrin_{34}$ are composed of 17 and 34 amino acid residues, respectively. $Gastrin_{17}$ is the most active and has a half-life of about eight minutes. Its precursor, $gastrin_{34}$, has a half-life of approximately 40 minutes. The release of gastrin is stimulated by food entering the stomach from the oesophagus and its function is, in turn, to stimulate the release of gastric juice. Gastric inhibitory peptide (GIP) is a peptide of 43 amino acid residues secreted by the duodenum and upper jejunum. It stimulates insulin release (*Chapter 7*), reduces the secretions of gastrin and pepsin, and inhibits gastric movements.

The hormone vasoactive intestinal peptide (VIP) comprises 28 amino acid residues. It is released in response to distension of the GIT by food. It stimulates the contraction of smooth muscle tissues of the GIT wall and pancreatic exocrine secretions, and it also inhibits gastrin and gastric acid release.

Secretin is also produced in the duodenum and jejunum. It is a 27 amino acid residue peptide with a half-life of 17 minutes. Its release is triggered by stomach acid and stimulates the release of pancreatic juice, the HCO_3^- of which helps neutralise the acid in chyme. Glucose-dependent insulinotrophic polypeptide (gastric inhibitory polypeptide) is also released in the duodenum and jejunum. It inhibits the secretion of gastric acid and stimulates insulin secretion. Mucosal cells in the upper region of the small intestine secrete cholecystokinin (CCK). Two molecular forms are produced, consisting of 33 and 39 amino acid residues, respectively. The release of CCK is stimulated by peptides and fatty acids in the food and, in turn, stimulates the release of pancreatic juice and contractions of the gallbladder. Motilin is a 22 amino acid residue peptide that is structurally unrelated to any other GIT hormone produced in the upper small intestine. It controls GIT movements during fasting.

The ileum and colon produce peptide YY and neurotensin. The former decreases pancreatic and gastric secretions, while the latter may regulate peristalsis of the ileum. The hormone called substance P is produced along the entire GIT. Its functions include stimulating the secretion of saliva and it is also involved in the vomit reflex.

Pancreatic polypeptide is formed, as its name implies, by the pancreas and inhibits pancreatic HCO_3^- and protein secretions.

11.10 DISORDERS OF THE GASTROINTESTINAL TRACT

Disorders of the GIT and its accessory organs can affect the mouth and oesophagus, stomach, small and large intestines, pancreas, liver, bile duct and gallbladder. Several of

these disorders also affect exocrine functions of the stomach, small and large intestines, pancreas and liver.

Disorders of the mouth

Xerostomia, an abnormally dry mouth, affects people who breathe through the mouth or produce insufficient saliva. Its incidence increases with age. It can also occur as a side effect of many medications. The lack of lubrication can impair swallowing food (dysphagia), which increases the risk of aspirating food into the trachea. The cracking and splitting of oral tissues also increases the risk of infection.

Infections of the mouth include oral herpes, commonly called cold sores. These infections with herpes simplex virus 1 (*Chapters 2 and 3*) generally affect the lips and around the mouth. Although often painful and very infectious, they normally heal within 7–14 days. In the early stages, they can be treated with acyclovir or valacyclovir (see *Figure 3.3*), but this is ineffective if a secondary bacterial infection has taken hold.

Mouth ulcers or canker sores are yellow-coloured areas with reddish borders on soft tissues in the mouth. Aphthous and traumatic ulcers are the most common types. The former is associated with stress or underlying illnesses and may also have a genetic contribution. Traumatic types are caused by injuries arising from overly hot foods, biting the inside of the mouth or other types of mechanical damage. Acute ulceration can be caused by viral or bacterial infections. While mouth ulcers are painful, most heal spontaneously within 10 days.

Infections with yeasts (thrush), such as *Candida albicans*, can cover all areas of the mouth. This is generally a sign of stress or immunodeficiency, for example human immunodeficiency virus (HIV) infection (*Chapters 2* and *3*) or chemotherapy to treat cancer (*Chapter 17*). There is evidence that, in the UK, cancers of the mouth are increasing.

Gingivitis or infection of the gums is a common condition resulting in swollen, sore gums that can bleed. It normally results from poor oral hygiene, leading to an accumulation of dental plaque, that is, a thin layer of mucins combined with salivary proteins and often oral bacteria. If not removed by appropriate brushing, plaque can develop into tartar or dental calculus, which consists of a firmly adhered layer of calcium salts and organic matter that is more difficult to remove (*Margin Note 11.6*). However, diabetes, hormonal changes associated with pregnancy or clinical conditions that impair the immune system can all result in plaque deposition.

Untreated gingivitis can result in periodontitis, where the tissues that support the teeth become infected. Its symptoms can include bad breath (halitosis), an unpleasant taste in the mouth, loose teeth that can make eating difficult, bleeding, painful gums and gum abscesses with the accumulation of pus in the gums and under the teeth. A recurrent periodontal disease of sudden onset called acute necrotising ulcerative gingivitis (ANUG) can develop, with usually more severe symptoms, including bone destruction and fever. If not removed, bacteria in plaque and tartar produce acids and enzymes that can demineralise teeth and degrade their protein content. This can lead to the formation of dental cavities in teeth and result in dental caries where bacteria infect the pulp.

Treatment for mouth disorders depends on the condition and its cause. If it is caused by underlying disease, then that must be addressed. Mild cases of gum disease can usually be treated by increased oral hygiene following professional oral hygiene advice, regular dental inspections and not smoking. Treatments for periodontitis and ANUG may include proprietary painkillers, mouthwashes containing chlorhexidine or hexetidine, and possibly the antibiotics metronidazole or amoxicillin. In severe cases, periodontal surgery and tooth removal may be necessary. The treatment of dental cavities and removal of decayed teeth should only be effected by professional dentists.

Margin Note 11.6
Dental hygiene

The minimum standard of oral hygiene consists of the following:

- twice daily brushing of teeth for two minutes
- regular flossing and/or cleaning between teeth with interdental brushes
- four thorough cleanings of teeth annually by a dental hygienist to remove firmly adhered plaque and tartar
- twice yearly dental checks by a trained dentist to detect and treat any damage to teeth

Disorders of the oesophagus and stomach

The following disorders of the stomach will be considered:

- hiatus hernias
- gastritis
- ulcers
- Zollinger–Ellison syndrome

Hiatus hernias

Hernias are protrusions of a tissue or organ out of the body cavity in which it is normally found. A hiatus hernia (*Figure 11.37*) occurs when a defective diaphragm allows the upper part of the stomach to enter the thoracic cavity through the oesophageal hole of the diaphragm. Sliding hiatus hernias occur when the oesophagus and stomach both move upwards so that the top end of the stomach protrudes through the gap in the diaphragm

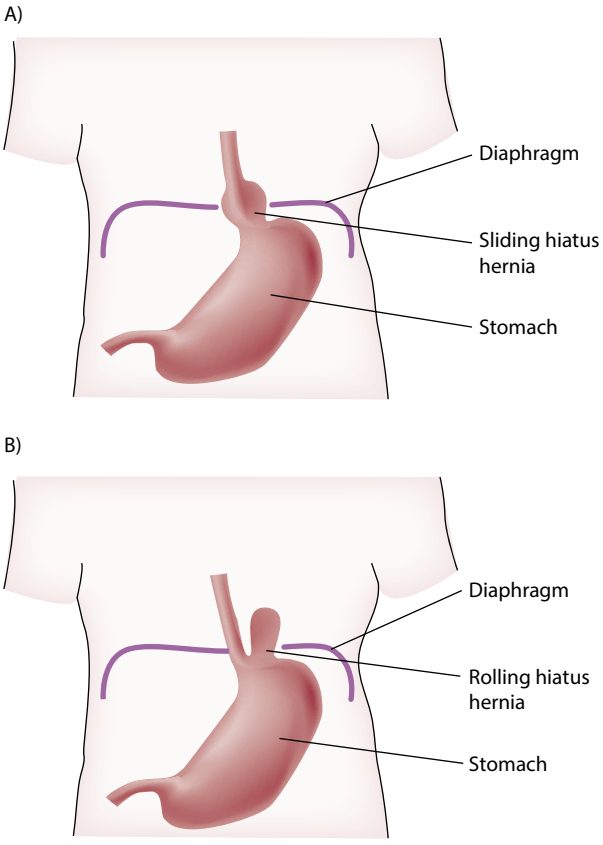


Figure 11.37 Schematics showing (A) sliding and (B) rolling hiatus hernias.

normally occupied by the oesophagus and these constitute 90% of cases. The remaining 10% are rolling hiatus hernias where a portion of the stomach curls upwards adjacent to the oesophagus so that both it and an upper part of the stomach protrude through the gap (see *Figure 11.37*). Data on the incidence of hiatus hernia are few but the condition increases with age and is particularly common in overweight middle-aged women and can also occur during pregnancy. The causes of hiatus hernias are unknown but they may be due to intra-abdominal pressure or weakening of the gastro-oesophageal junction muscles caused by trauma or loss of tone. Over 50% of individuals with hiatus hernia are asymptomatic, but when symptoms do occur, they include heartburn that is aggravated by reclining, chest pain, dysphagia, belching, pain on swallowing hot fluids and a feeling of food sticking in the oesophagus. Although hiatus hernias are not usually serious, they can cause inflammation of the oesophagus leading to a back flow of gastric juices called **reflux oesophagitis**, and it may cause bleeding (perhaps leading to anaemia) or a stricture (narrowing) of the oesophagus, and acid burns to the nasopharyngeal region. Cancers associated with hiatus hernias are very rare, but there is a slightly increased risk of them developing in inflamed areas of the oesophagus.

A preliminary diagnosis requires confirmation by gastroscopy in which a tube with a camera at its end is passed into the stomach, allowing a direct visualisation of its wall or by X-ray analysis. However, the contents of the GIT are often not clearly visible using X-rays, and diagnosis may require ingesting a barium meal consisting of barium sulphate mixed with a, usually flavoured, liquid. The barium salt lines the inside of the GIT wall and is visible because barium is opaque to X-rays making this a useful method for detecting structural abnormalities of the GIT. The aims of treating hiatus hernias are to alleviate their symptoms. Losing weight, reducing smoking, coffee and alcohol intakes, and avoiding tight or restrictive clothing all help to relieve symptoms. Elevating the head of the bed when sleeping can help to reduce acid reflux. Medication such as antacids or H^+ pump inhibitors, for example lansoprazole or esomeprazole, or a histamine-2 receptor antagonist such as ranitidine may be prescribed. Surgery is only used when there is strangulation of the hernia or the symptoms cannot be controlled.

A condition that can resemble symptomatic hiatus hernias is gastro-oesophageal reflux disease (GORD). Indeed, symptomatic sufferers of hiatus hernias frequently experience GORD, which causes frequent regurgitation of stomach and duodenal contents to occur. Episodes generally happen at night. Mild episodes are often described as heartburn, but serious ones result in acid damage to the nasopharyngeal region, prolonged coughing,

vomiting and possibly damage to teeth. Persistent acid damage to the lower oesophagus can result in Barrett oesophagus, a precancerous condition where the lining resembles that of the upper stomach. Its development to a cancerous condition is rare, with oesophageal cancers constituting 3% of UK cases (*Chapter 17*).

Gastritis

Gastritis is the most common disorder affecting the stomach. It is characterised by inflammation and erosion of the gastric mucosa. Gastritis is idiopathic in many cases, but it can be caused by irritating foods and beverages, ingested poisons, aspirin and staphylococcal exotoxin (*Chapters 2* and *3*). Gastritis may present **acutely**, where the patient suffers GIT bleeding, epigastric pain (that is, pain on or over the stomach area), anorexia and vomiting of blood (**haematemesis**). Patients with **chronic** gastritis may lack symptoms except for epigastric pain. The possibility of exposure to irritating substances must be determined when assessing the clinical history of the patient. Chronic gastritis frequently is asymptomatic but can present as nonspecific abdominal pain. Since gastritis often occurs in severely ill, hospitalised people, its symptoms may be eclipsed by other, more severe symptoms. Gastroscopy can confirm a diagnosis of gastritis by revealing inflamed portions of the lining of the stomach. Relief from the symptoms of gastritis occurs following removal of the irritant substance or treatment of the underlying cause(s).

Atrophic gastritis is a degenerative stomach disorder characterised by chronic inflammation of the stomach with atrophy of its mucous membrane lining (*Figure 11.38*). Specific data on the incidence of atrophic gastritis are scarce. However, its prevalence mimics that of its two main causes. It can be caused by persistent infection with the bacterium *Helicobacter pylori*, but it can also have an autoimmune origin. In both types, atrophic gastritis develops over many years and is detected later in life. *H. pylori* (*Figure 11.39*) infects approximately 20% of people younger than 40 years and 50% of those older than 60 years in the developed world. Infection is highly prevalent in Asia and in developing countries and it is estimated that 50% of the world's population is infected. Thus, chronic gastritis is probably extremely common but is usually detected late in life. The condition results in loss of gastric glandular cells with their eventual replacement by nonsecretory and fibrous tissues. Secretions of HCl, pepsin and intrinsic factor are impaired, leading to digestive problems, vitamin B_{12} deficiency and megaloblastic anaemia.

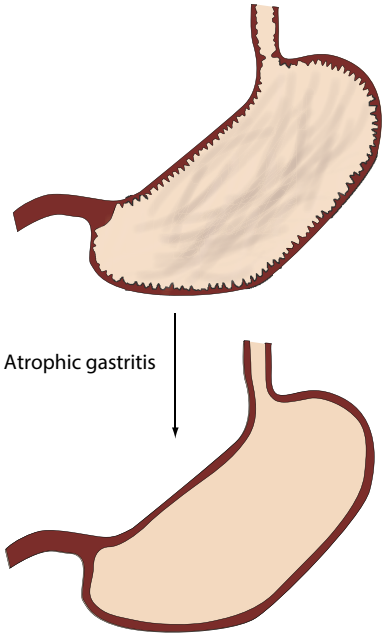


Figure 11.38 Schematic showing the effects of atrophic gastritis on the stomach lining. ⏎

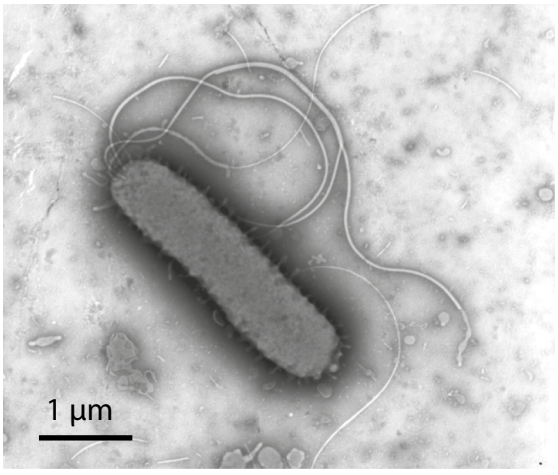


Figure 11.39 Electron micrograph of *Helicobacter pylori*. Courtesy of Drs A. Curry and D. Jones, Public Health Laboratory Service, Withington Hospital, UK. ⏎

H. pylori can bind to the stomach lining. It expresses urease, which hydrolyses urea releasing ammonia (see *Figure 11.41*) that neutralises stomach acid allowing the bacterium to penetrate into the mucosal layer. The release of bacterial and inflammatory toxic products by *H. pylori* over time results in increasing gastric mucosal atrophy. Some glandular units develop an intestinal-type epithelium; others are simply replaced by fibrous tissue. The loss of gastric mucosa decreases HCl secretion, which increases the gastric pH and leads to a reduced ability to kill bacteria. Ingested bacteria can survive and reside in the stomach and the upper part of the small intestine. Infection is usually acquired during childhood and, if left untreated, progresses over the lifespan of the individual in one of two main ways that have different pathological consequences. The first is a gastritis that mainly affects the antrum of the stomach. This is the most frequently observed pattern in Western countries, and individuals with peptic ulcers (see below) usually followed this clinical history. The second pattern is a more widespread atrophic

gastritis affecting, for example, the corpus, fundus and antrum with the loss of gastric glands and their partial replacement by an intestinal-type epithelium. This pattern is observed more often in developing countries and Asian individuals who develop gastric carcinoma and gastric ulcers and usually present with this pattern of gastritis.

Atrophic gastritis cannot be reliably diagnosed by gastroscopy but requires a microscopic examination of biopsy specimens. *H. pylori* infections are normally diagnosed using serological tests, breath tests or antigen tests of the faeces. Pernicious anaemia resulting from autoimmune atrophic gastritis usually presents in patients approximately 60 years of age. Treatment of atrophic gastritis is directed at eliminating the causative agent, correcting complications of the disease and attempting to revert the atrophic process. When *H. pylori* is the causative agent, it can be eradicated using a combination of antimicrobial agents and antisecretory agents with a success rate of about 90%. Lack of patient compliance and antimicrobial resistance are the most important factors influencing poor outcomes. However, treatment of *H. pylori* infection may not lead to a reversal of existing damage unless started early, but can block further progression of the disease. Some evidence suggests that β carotene and/or vitamins C and E may reverse or reduce the risk of atrophic gastritis and/or gastric cancer.

Autoimmune gastritis is associated with serum anti-intrinsic factor antibodies that reduce the amount of functioning GIF. The prevalence of pernicious anaemia resulting from autoimmune gastritis is estimated to be 0.1% of the general UK population, increasing to 1.99% in those over 60 years old. Cell-mediated immunity also contributes to the condition because T cell lymphocytes infiltrate the gastric mucosa and contribute to the epithelial cell destruction and resulting gastric atrophy. Autoimmune gastritis is a relatively rare condition, which is most frequently observed in patients of northern European descent and in African Americans. If untreated, the major complication in autoimmune gastritis patients is the development of pernicious anaemia because of the decreased availability of vitamin B_{12} (*Chapter 13*). This requires vitamin B_{12} replacement therapy.

Ulcers

Ulcers are perforations of the GIT wall, particularly erosions of the mucosal layer related to cancer, that is, malignant ulcers, or to stomach HCl, that is, peptic ulcers. Ulcers may also be named from their location, for example mouth, oesophageal, gastric or stomach, and duodenal ulcers. Oesophageal ulcers are usually associated with hiatus hernias (see above) caused by acid splashing from the stomach (acid refluxing) into the oesophagus. Gastric ulcers are relatively rare because the mucosal lining of the stomach is protected from the acid by the layer of alkaline mucus. They generally occur in patients older than 50 years of age. It was thought that ulcers were caused by stress and excessive accumulation of HCl. However, it is now accepted that their most common cause is infection with *H. pylori* (see *Figure 11.39*), which can colonise and destroy the mucosal layer. Duodenal ulcers are five times more common than gastric ones and generally occur in a younger population. More than 90% of ulcers occur in the duodenal wall, usually after it has been weakened by infection with *H. pylori*.

Peptic ulcers are linked to an increased production of HCl and pepsin in gastric juice or to a reduced protection of the mucosa against gastric juice. *Figure 11.40* illustrates the appearance and (diagrammatically) the development of a peptic ulcer. Lesions that do not extend through the mucosal lining are referred to as erosions. Acute and chronic ulcers erode this layer and, in serious cases, may penetrate through the stomach wall. In some patients, blood vessels in the GIT wall ulcerate and lead to heavy, and in some cases fatal, bleeding. Chronic ulcers have an associated basal scarring.

Patients with peptic ulcers present with heartburn, indigestion, gastrointestinal bleeding, vomiting, weight loss and upper abdominal pain that can be relieved by foods and antacids. A diagnosis is made on clinical grounds supported by endoscopy, but laboratory tests that can assist include the urea breath test, the *H. pylori* immunoglobulin G (IgG) antibody test and the faecal *H. pylori* antigen test. In the urea breath test, the patient is given a fruit drink containing ^{13}C- or ^{14}C-labelled urea. Urease activity by *H. pylori* converts the urea to $^{13/14}CO_2$ and NH_3. A breath sample is taken after 30 minutes and its $^{13/14}CO_2$ content measured. A high exhaled $^{13/14}CO_2$ indicates the presence of *H. pylori* in the GIT (*Figure 11.41*). The *H. pylori* IgG antibody test identifies antibodies against the bacterium and is used to establish if an infection is present. It is not used to monitor treatment, as residual antibodies are present after treatment and will be detectable. The faecal *H. pylori* antigen test directly measures antigens derived from the pathogen in the faeces of possible patients. The test is used to establish an infection and to monitor its treatment.

The treatment of peptic ulcers is aimed at eradicating the *H. pylori* infection and reducing acid output. Antibiotics (*Chapter 3*) that effectively suppress symptoms include amoxycillin, clarithromycin, metronidazole and tetracycline, and they often cure the

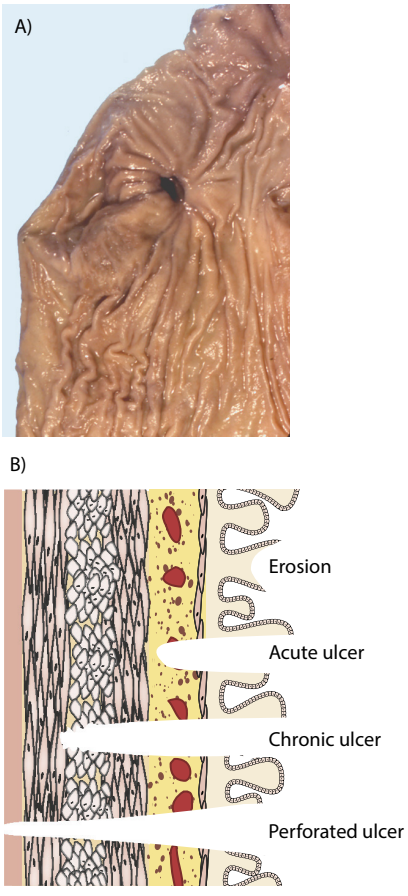


Figure 11.40 (A) Picture of a gastric ulcer. (B) Schematic showing the development of an ulcer in the stomach wall. (A), courtesy of Dr A.S. Mills, Virginia Commonwealth University, USA. ⏎

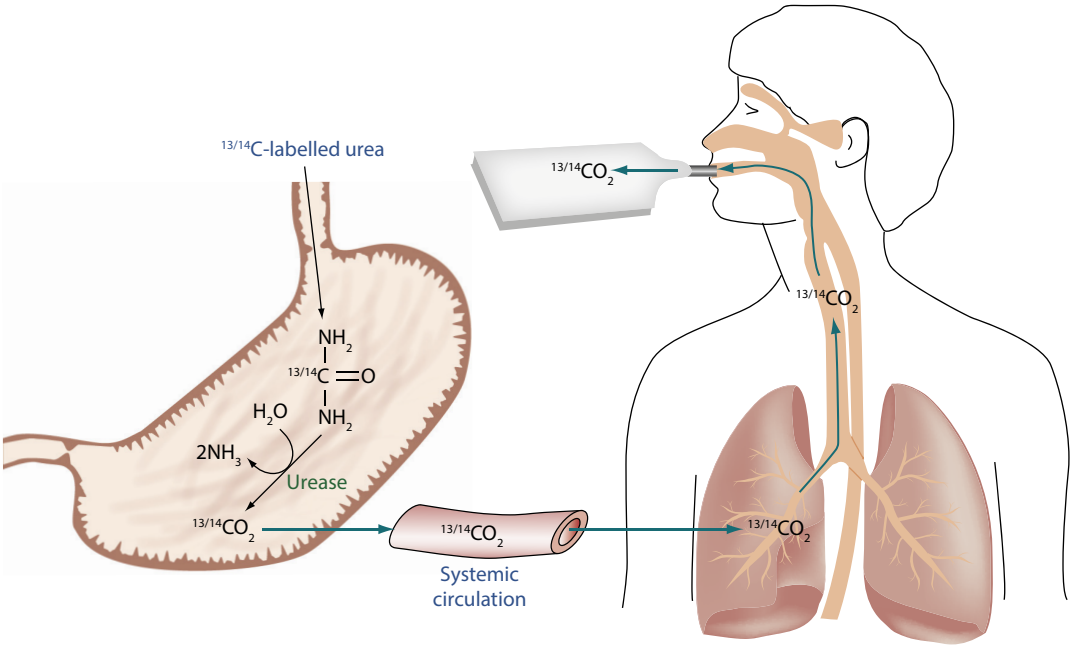


Figure 11.41 Overview of the urea breath test. Blue arrows represent movement of the $^{13/13}CO_2$.

patient. Bismuth chelate and sucralphate may also be administered to decrease the biosynthesis of prostaglandins that stimulate inflammation. The resulting decrease in acid production by parietal cells and the increase in HCO_3^- production by mucus-secreting epithelial cells have cytoprotective effects.

Zollinger–Ellison syndrome

Zollinger–Ellison syndrome causes massive, multiple and recurrent peptic ulcers due to the excessive secretion of gastric juice from tumours affecting the pancreas or duodenum. It is a rare disorder affecting about 1 in a 1 000 000, but approximately 60% of the tumours are malignant and are referred to as gastrinomas because they secrete large amounts of gastrin. The concentration of gastrin in the plasma is frequently increased to values above 200 ng/L (normal values are less than 100 ng/L). However, in 30% of patients, the concentration of gastrin in the plasma is normal or only slightly raised. The syndrome presents with abdominal pain, nausea, vomiting, bleeding in the stomach, diarrhoea, weakness and fatigue. Many patients also suffer from fat malabsorption because the high acidity reduces activity of pancreatic lipase. Potential patients are investigated by measuring the secretion of HCl. A preliminary diagnosis is confirmed if this is greater than 100 mmol/hour (normal rates are less than 5 mmol/hour). Investigations of patients also include endoscopy to detect ulceration. Patients with Zollinger–Ellison syndrome are treated using H^+ pump inhibitors to reduce acid secretion and surgery to remove the gastrinoma; radiotherapy and chemotherapy may be used when malignant tumours are present (*Chapter 17*).

Disorders of the small intestine

Disorders of the small intestine include:

- lactose intolerance
- coeliac disease
- dietary malabsorption

Lactose intolerance

Lactose intolerance arises from an inability to express lactase. It is divided into three categories:

1. congenital alactasia
2. primary acquired lactose intolerance
3. secondary acquired lactose intolerance

Congenital alactasia or hypolactasia is extremely rare. Affected babies do not gain weight, are dehydrated and are extremely unwell. Human milk is unsuitable for affected

babies and breastfeeding is precluded, which can also cause emotional distress to some mothers. These babies must be fed lactose-free dairy- or soya-based formulae to survive. Primary acquired lactose intolerance usually occurs following weaning and before the age of six years, and is the normal condition for approximately 70% of the world's population, the major exception being northern Europeans. It is particularly common in Asian communities and amongst Black people of African origin. Secondary acquired lactose intolerance occurs because of damage to the small intestinal mucosa, for example due to gastroenteritis, cow's milk protein intolerance or coeliac disease (see below).

Patients who ingest milk present with serious indigestion, nausea, abdominal distension, diarrhoea and increased flatus because of the action of GIT bacteria on the ingested lactose. The severity of the symptoms vary with the amount of lactose consumed and the tolerance level of the individual. These rather diffuse symptoms are often associated with other conditions of the GIT, such as infections with parasitic helminths and the protozoan *Giardia* (*Chapters 2* and *3*); inflammatory conditions, for example ulcerative colitis (see below); hormonal complaints, such as hypo- and hyperthyroidism (*Chapter 7*); and cancers of the colon and rectum (*Chapter 17*).

Lactose intolerance can be investigated using the stool acidity test. Its principle is simple: undigested lactose fermented by bacteria in the colon produces lactate and fatty acids, which can then be detected in a stool sample. If this test is inconclusive, patients with suspected lactose intolerance are investigated using the more invasive lactose tolerance test, which essentially measures the malabsorption of lactose. The patient is required to fast overnight, then given a 50 g lactose meal in solution and the concentration of glucose in the blood is measured at baseline and regular intervals over the next three hours. A failure to increase the concentration of glucose by 1.1 mmol/L above the baseline value over this period suggests lactose malabsorption due to deficiency of lactase. The medical history, age, degree of intolerance and overall health of the patient determine the treatment for lactose intolerance. In general, patients are managed by limiting the intake of lactose-containing foods. Lactose intolerant patients can ingest a variety of milk products, such as cheese, where the lactose has been removed in the whey, and yoghurt, where it has been fermented to lactate. Adding proprietary products containing lactase (*Figure 11.42*) isolated from microorganisms to dairy products prior to ingestion can also help overcome the problem. However, patients on strict diets may require calcium and vitamin D supplements if their intakes are inadequate.

Figure 11.42 A proprietary product containing lactase used to digest lactose prior to ingestion.

Coeliac disease

Coeliac disease is a genetically determined chronic inflammatory condition affecting the small intestine. It is a relatively common disorder with a worldwide prevalence of 1 in 200–300. Coeliac disease is induced mainly by the ingestion of the wheat protein, gluten and its products. A portion of the gluten molecule forms an autoimmune complex in the GIT mucosa, which stimulates T cells to aggregate and release toxins that promote lysis of enterocytes (*Chapters 4* and *5*). This leads to a progressive atrophy and characteristic flattening of the mucosal villi and microvilli in the upper part of the small intestine (*Figure 11.43*). Coeliac disease is accompanied by malabsorption of nutrients due to a decreased surface area over which absorption can occur. The signs and symptoms of coeliac disease are surprisingly diverse and include abdominal pain, chronic diarrhoea, weight loss, bone pain, fatigue and anaemia. Coeliac disease is diagnosed by demonstrating villous atrophy in a small intestine biopsy (*Figure 11.44*) and by improvements in clinical symptoms or histological tests following restriction to a gluten-free diet. A demonstration of raised antibodies to tissue transglutaminase is also of utility in diagnosing the condition.

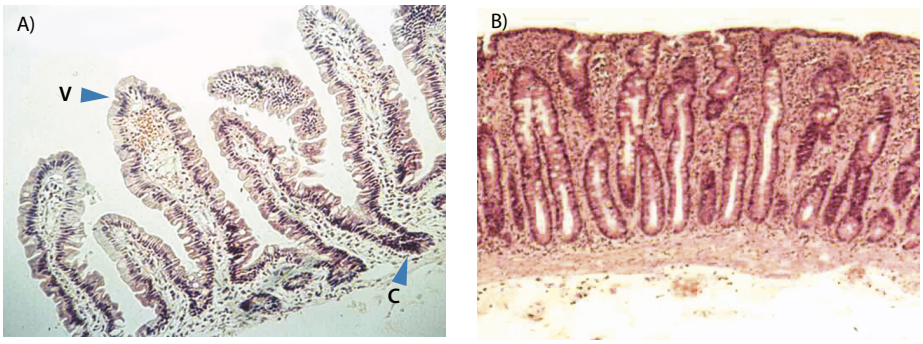


Figure 11.43 Photomicrographs of (A) healthy villi (V) and associated crypts of Lieberkühn (C) and (B) the mucosal lining from a coeliac patient. Note the degeneration of the villi.

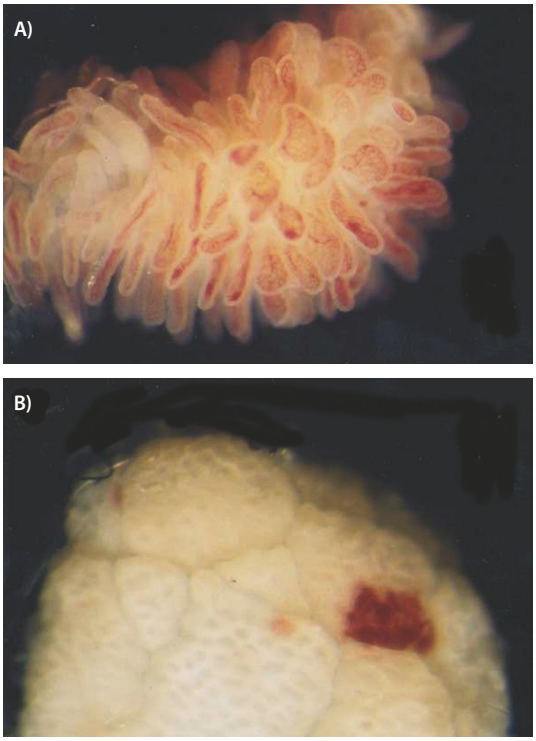


Figure 11.44 Photographs of gastrointestinal tract biopsies from (A) a healthy person and (B) a coeliac patient. Note the degeneration of the villi.

The management of coeliac disease involves adherence to a gluten-free diet. Gluten-free foods are readily available in supermarkets and health food stores, which in most cases helps to relieve the symptoms and allows the existing mucosal damage to heal, as well as preventing further damage. Relapses occur if gluten is reintroduced in the diet.

Dietary malabsorption

Dietary malabsorption is a reduced absorption of one or more nutrients by the small intestine. It occurs as a result of a wide range of disorders that affect the GIT, pancreas, liver and gallbladder. Its causes include enzyme deficiencies, such as in lactose intolerance; chronic pancreatitis; bile salt deficiency, as in biliary obstruction or hepatitis (see later); and intestinal diseases such as coeliac disease and Crohn disease (see below). The clinical features of malabsorption (*Figure 11.45*) arise because of deficiency of one or more nutrients and include anaemia due to iron, vitamin B_9 (folate) and/or vitamin B_{12} (cobalamin) deficiencies, osteomalacia where vitamin D is deficient, oedema caused by hypoalbuminaemia, the tendency to bleeding when vitamin K is suboptimal and generalised weight loss. Malabsorption results in retention of nutrients in the GIT lumen causing diarrhoea and abdominal discomfort following GIT bacterial metabolism.

The malabsorption of fats leads to their losses in amounts greater than 5 g daily and causes steatorrhoea: the faeces are greasy, with a pale colour and an offensive smell. A diseased liver may be unable to synthesise bile, which is required for digestion of fats and if absent can lead to malabsorption. Bacteria colonising the small intestine may break down bile acids, reducing their effective concentrations and leading to fat malabsorption. Malabsorption of fat also occurs in several intestinal and pancreatic disorders.

A faecal fat test can assess fat malabsorption. The test involves collecting faeces over a period of three days after which their fat content is assessed chemically. Normally, up to 5 g of fat is lost in the faeces each day but more is lost during malabsorption.

Liver function tests (LFTs; see *Box 11.10*) can indicate the presence of liver disease, and a number of blood tests are also useful in assessing liver function. These include determining the concentration of albumin in the plasma and haematological investigations such as full blood count, and vitamin B_9 (folate), vitamin B_{12} (cobalamin) and iron determinations can indicate the type of malabsorption. Investigations of malabsorption also include microbiological examination of faeces to identify any pathogen present.

The pentose sugar xylose is absorbed in the small intestine but is not metabolised and is excreted unchanged in the urine. This property is exploited in the xylose absorption test to assess the malabsorption of carbohydrates. The patient fasts overnight, empties the bladder and drinks a 500 mL solution containing 5 g xylose. In normal individuals, the serum xylose concentration increases above 1.3 mmol/L one hour after ingestion.

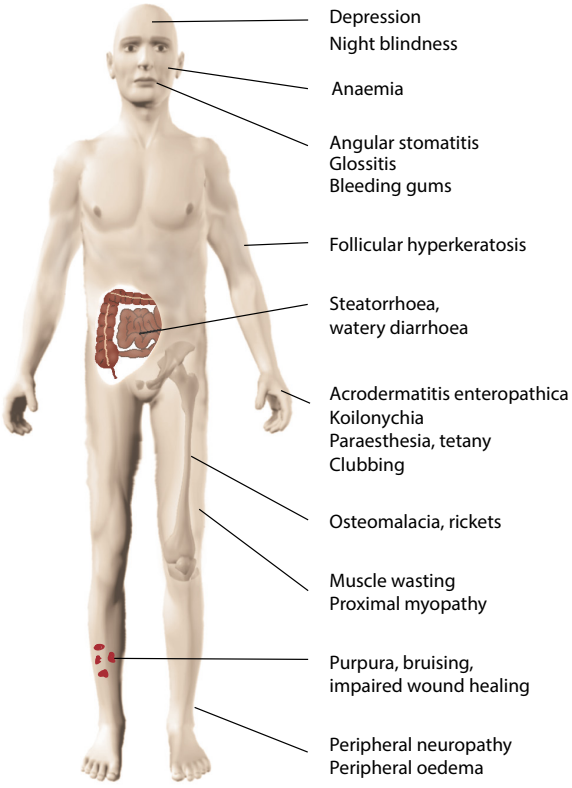


Figure 11.45 Schematic showing the clinical features associated with malabsorption.

BOX 11.7 HARTNUP DISEASE

Hartnup disease, also known as Hartnup disorder, Hartnup aminoaciduria or Hartnup syndrome, was first named by Denis Baron and coworkers in 1956 from an amino acid absorption disorder that affected the Hartnup family of London. It has an overall prevalence of 1 in 18 000–42 000 and, although rare, this makes it one of the most common amino acid disorders. Hartnup disease is inherited as an autosomal recessive trait (*Chapter 15*), particularly where consanguinity is common. It arises from mutations to *SLC6A19*, a gene located on chromosome 5, whose product, the Na^+-dependent neutral amino acid transporter B(0)AT1, is expressed mainly in the GIT and kidneys. The defective gene product impairs the absorption of tryptophan and other neutral amino acids, such as valine, phenylalanine, leucine and isoleucine, across the brush border membranes of the small intestine and renal tubular epithelium. Two tissue-specific forms have been described; one affects both the GIT and the kidneys, the other only the kidneys. The abnormality in amino acid transport can lead to deficiencies in neutral amino acids and the defective absorption of tryptophan and so a vitamin B_3 (niacin) deficiency (*Chapter 10*). The condition clinically resembles pellagra and so a misdiagnosis of the vitamin deficiency is possible. Tryptophan is retained within the GIT lumen and can be converted to toxic indole compounds by bacteria. Tubular renal transport is also defective and contributes to gross aminoaciduria.

Hartnup disease usually begins at three to nine years of age, but it may present as early as 10 days after birth. Most patients are asymptomatic, but poor nutrition leads to more frequent and severe attacks. Patients present with pellagra-like light-sensitive rash, aminoaciduria, cerebellar ataxia, emotional instability, and neurological and psychiatric symptoms that may considerably diminish their quality of life. Mental abilities are diminished and short stature have been described in some patients. Malnutrition and a low-protein diet are the primary contributing factors to morbidity; in rare cases, severe central nervous system (CNS) damage may lead to death.

Plasma concentrations of amino acids are usually normal in Hartnup disease. However, it may be diagnosed by demonstrating increased concentrations of neutral amino acids, such as tryptophan, valine, phenylalanine, leucine and isoleucine in the urine, with normal levels of proline, hydroxyproline and arginine, which differentiates Hartnup disease from other causes of gross aminoaciduria. Other tryptophan metabolites, including kynurenine and serotonin, are also found in the urine, and neutral amino acids occur in the faeces. Urinary indole derivatives, for example 5-hydroxyindoleacetic acid, may be demonstrated following an oral tryptophan load (see *Figure 10.45*). In selected patients, diagnosis may require the transport defect to be identified *in vitro* using a jejunal sample obtained by biopsy.

Most patients can be treated with a high-protein diet, which overcomes the deficient transport of neutral amino acids because dipeptides can be actively absorbed, as shown in *Figure 11.25*, and by paracellular routes. Indeed, in the developed world, cases presenting with extreme clinical symptoms are rare, probably because of the protein-rich diet available. Patients who are symptomatic are advised to avoid excessive exposure to sunlight and use sunscreens with a skin protection factor of at least 15. Dietary supplements of 40–200 mg niacin daily or nicotinamide or tryptophan ethyl ester may relieve the pellagra-like symptoms. Psychiatric treatment may help patients with severe CNS damage.

After five hours, the concentration of xylose in the urine increases to more than 7.0 mmol/L. Significantly lower concentrations of xylose occur in the serum of patients with carbohydrate malabsorption. However, care must be exercised, since some bacteria colonising the small intestine are capable of metabolising xylose, and several renal diseases can also lead to reduced concentrations.

Disorders of the large intestine

Disorders of the large intestine include:

- Crohn disease
- ulcerative colitis
- haemorrhoids
- irritable bowel syndrome
- short bowel syndrome

Crohn disease

Crohn disease, named after the American physician Burrill Crohn (1884–1983), is a chronic inflammatory condition of the bowel that affects the terminal ileum, jejunum and colon. The inflammation extends throughout the layers of the intestinal wall, thickening it and narrowing the lumen, and can cause obstruction of the intestine. Diseased areas tend to be patchy, being separated by normal bowel segments giving characteristic so-called 'skip lesions'. The cause of Crohn disease is unclear, although viruses and bacteria have been implicated. Patients with the disease suffer from a lack of appetite, abdominal pain, weight loss, fatigue, diarrhoea and bloody stools, and malabsorption.

Crohn disease is investigated by determining the presence of faecal calprotectin, a metalloprotein (*Figure 11.46*) released into the small intestine during inflammation. Its measurement is useful for distinguishing between inflammatory conditions, for example Crohn disease and ulcerative colitis, and noninflammatory bowel disease, such as irritable bowel syndrome (see below). A biopsy of the GIT may also be used to detect the characteristic changes associated with Crohn disease. Treatment involves using anti-inflammatory drugs, such as 5-aminosalicylic acid (*Figure 11.47*), although surgery may be required in severe cases.

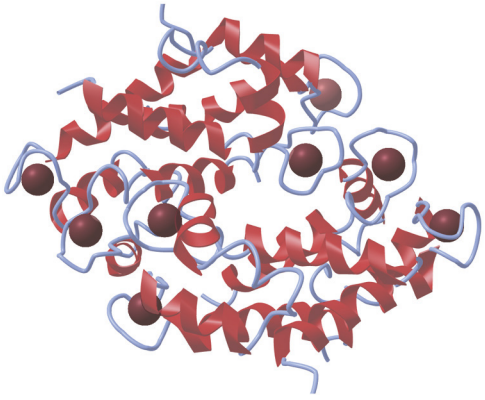


Figure 11.46 Molecular model of calprotectin. PDB file 1XK4.

Ulcerative colitis

Ulcerative colitis is characterised by chronic inflammation and ulceration of the colon that is distinctive in that it affects only its mucosal lining. The lining is also affected by numerous haemorrhagic ulcerations. The cause of ulcerative colitis is not entirely clear but is believed to be autoimmune in origin (*Chapter 5*). Patients with ulcerative colitis suffer from abdominal pain, fever, weight loss and bloody diarrhoea. Diagnosis of ulcerative colitis is made following colonoscopy where a biopsy from the colon is taken to detect the characteristic histological changes. Ulcerative colitis is treated with anti-inflammatory drugs, such as 5-aminosalicylic acid (*Figure 11.47*), although surgical removal of the affected region may be necessary in severe cases.

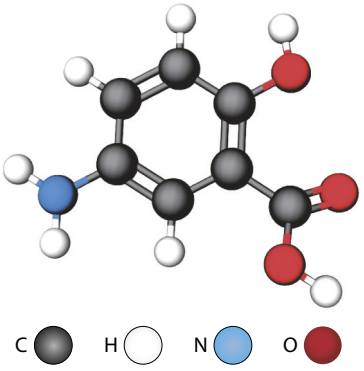


Figure 11.47 5-Aminosalicylic acid.

Irritable bowel syndrome

Irritable bowel syndrome (IBS) is characterised by abnormally excessive changes in the motility of the small and large intestines resulting in constipation or diarrhoea or alternations between the two. The syndrome is most common in 20 to 40 year olds, with the

BOX 11.8 DIARRHOEA

Diarrhoea occurs when is an imbalance in the two-way flow of water and electrolytes, which occurs between the gastrointestinal tract (GIT) lumen and the basolateral extracellular fluid (ECF). Water enters the GIT in food and drinks and in various secretions. In the small intestine, the secretion of water and electrolytes normally occurs in the crypts of Lieberkühn. Sodium chloride is transported from ECF into epithelial cells across their basolateral membranes. Secretory stimuli increase the permeability of luminal membranes of the crypt cells, allowing Cl^- to move into the GIT lumen. Sodium ions, however, are returned to the ECF by the action of the Na^+/K^+-ATPase of the basolateral membranes. The movements of these ions generate an osmotic gradient and water flows passively from the ECF into the lumen through intercellular channels. Water reabsorption is also driven by osmotic gradients, which are formed when solutes, particularly Na^+, actively enter enterocytes. Absorption of Na^+ may occur by direct transport of the ion or be exchanged for H^+ or linked to the absorption of glucose, amino acids or Cl^-. Following absorption, Na^+ is transported out of enterocytes across the basolateral membrane into the ECF by the Na^+/K^+-ATPase. This increases the osmolality of the ECF and water moves passively into it from the GIT lumen. These processes maintain an osmotic balance between GIT contents and ECF in the intestinal tissue. However, since fluid absorption normally exceeds secretion, the net result is fluid absorption. More than 90% of the fluid entering the small intestine is absorbed, with only about 1 L reaching the large intestine. Here, further absorption occurs and only 100–200 mL of water is lost in the faeces daily. Diarrhoea arises when water and electrolyte transport becomes disordered, for example by increased secretion, decreased absorption or both, and an increased volume of fluid enters the large intestine. When this exceeds the absorptive capacity of the large intestine, it results in diarrhoea.

Diarrhoea may result from one of two principal mechanisms: secretion and osmotic imbalance. These mechanisms are not exclusive; intestinal infections can cause both types and, indeed, both may occur in a single individual. Secretory diarrhoea is the more common of the two and is caused by an abnormal secretion of water and salts into the small intestine. This occurs when the reabsorption of Na^+ is impaired but the secretion of Cl^- in the crypts of Lieberkühn is maintained or increased. This produces a net secretion of fluid resulting in the loss of water and salts from the body in watery stools that leads to dehydration. Infectious diarrhoea may result from the actions of parasites, bacterial toxins or viruses on the GIT mucosa (*Chapter 2*). Osmotic diarrhoea results when a poorly absorbed, osmotically active substance is ingested, causing water and salts to move rapidly across the GIT lining to maintain the osmotic balance. The effects depend upon the osmolarity of the solution. If the substance is taken as an iso-osmotic solution, water and solute will pass through the GIT, with no net absorption causing diarrhoea. If taken as a hyperosmotic solution, water and some electrolytes will move from the ECF into the GIT lumen, increasing the volume of the faeces and causing dehydration. Furthermore, because the loss of body water is proportionally greater than the loss of Na^+ and Cl^-, hypernatraemia also develops (*Chapter 8*). Diarrhoea results in losses of large amounts of water, Na^+, K^+, Cl^- and HCO_3^-. Dehydration is the most clinically dangerous effect of diarrhoea in the short term, but **hypovolaemia** can result in cardiovascular collapse and death if not treated promptly (*Chapter 8*). The acute effects of electrolyte depletion include metabolic acidosis (*Figure 11.48*). The aims of managing diarrhoea are to correct the dehydration and electrolyte deficits. Fluids can be replaced orally or intravenously.

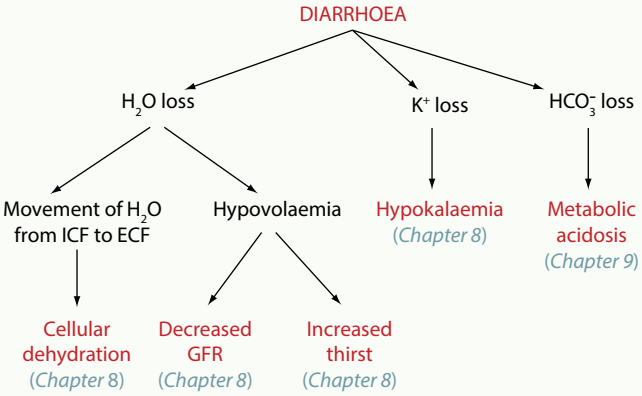


Figure 11.48 Examples of major clinical features that may arise from prolonged diarrhoea. ECF, extracellular fluid; GFR, glomerular filtration rate; ICF, intracellular fluid.

majority of sufferers being female. Symptoms include bloating, abdominal distension and chronic pain or discomfort in the lower abdomen; the latter is usually relieved by passing flatus (wind) or defaecation. The cause is unknown, but its development often follows an acute GIT infection, implicating an immune contribution. Diagnosis usually involves a clinical history, especially of bowel habits, and a physical examination of the abdomen to eliminate other clinical conditions that present with similar symptoms, such as lactose intolerance, dysentery, inflammatory bowel syndrome or cancers. Treatment is aimed at relieving the symptoms and includes the use of dietary soluble fibre agents that stabilise the faecal aqueous content and laxatives in constipation-dominant IBS. In contrast, antidiarrhoea agents are used in diarrhoea-dominant IBS. The use of mild tranquilisers and antidepressants can sometimes help if anxiety or depression are underlying contributory causes.

Short bowel/gut syndrome

Short bowel/gut syndrome refers to any malabsorption syndrome following surgical resectioning of the GIT (*Figure 11.49*), which has left an inadequate surface area for nutrient absorption. The extent of clinical problems depends upon the degree and type of

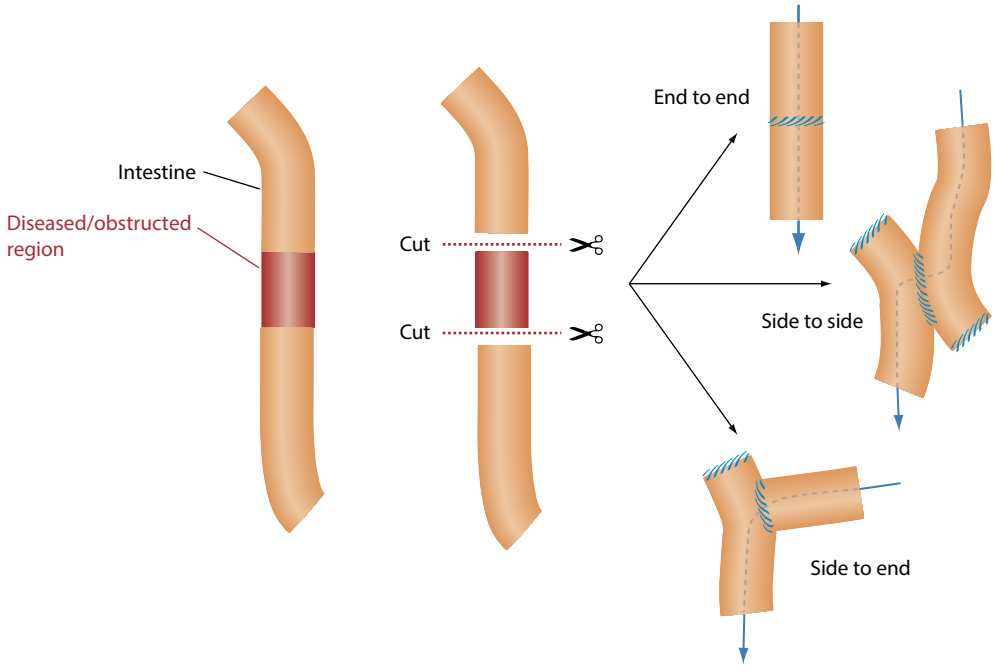


Figure 11.49 Schematic overview of types of surgical resectioning of the gastrointestinal tract.

resectioning, but symptoms may include malnutrition with loss of weight, weakness and dizziness, diarrhoea, the production of abnormal faeces and steatorrhoea. Diagnosis obviously includes taking a clinical history, particularly of the surgery involved. Treatment is dependent on the extent of the condition. Oral or intravenous food supplements can correct the malnutrition. In general, patients eventually recover even following extensive resection.

The condition can also occur in infants, where it results from an underdeveloped intestine or one that has yet to grow to normal dimensions. The child must be treated with parenteral feeding until the syndrome is corrected by natural growth and development of the intestine or by replacement surgery.

Haemorrhoids

Haemorrhoids are permanently dilated, varicose veins in the anal region. Internal haemorrhoids originate above the internal anal sphincter but can grow sufficiently large to protrude out of the anus (*Figure 11.50*). Straining to defaecate, constipation and sitting for extended periods can all lead to their development. In women, they are often associated with pregnancy. Severe constipation, a frequent but unrelieved urge to defaecate

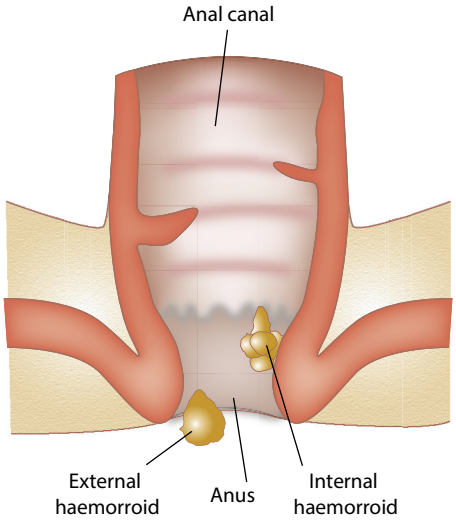


Figure 11.50 Haemorrhoids.

and bright red rectal bleeding are common symptoms. However, the patient should also be examined for other clinical conditions, for example bowel cancer, which present similar indications (*Chapter 17*). Haemorrhoids can be asymptomatic but, especially external ones, can be itchy, bleed and cause pain. Diagnosis is usually by visual inspection of the anal area, including proctoscopy to examine for internal haemorrhoids. Treatments include recommendations of high-fibre diets and dietary soluble fibre agents that soften the faeces, salt baths and topical hydrocortisone creams to reduce any inflammation. Haemorrhoids may be destroyed by band ligation. This simple and effective procedure does not require anaesthesia. The haemorrhoid is extended using forceps and a rubber band placed over it at its base. The constriction of blood supply results in tissue necrosis and sloughing of the material, usually within seven days. If the patient maintains a regular bowel habit and avoids constipation by ingesting a diet high in fibre and using soluble high fibre laxatives, while avoiding prolonged sitting and lifting heavy objects, complications and recurrences are uncommon.

11.11 DISORDERS OF THE GASTROINTESTINAL SYSTEM ACCESSORY ORGANS

As might be expected, clinical conditions and diseases affect all of the GIT accessory organs. In a text of this size, only a restricted number of these will be considered.

Disorders of the exocrine pancreas

Acute pancreatitis is a severe, rapid inflammation of the pancreas with varying degrees of oedema (*Chapter 8*), haemorrhage (*Chapter 13*) and tissue necrosis. It arises because of an inappropriate activation of pancreatic enzymes, which then autodigest pancreatic tissues. Normally, these enzymes are inactive until they reach the duodenum. The cause(s) of acute pancreatitis is/are unclear, although excessive alcohol intake is believed to play a major role, but viral infections, drug reactions and pancreatic cancer (*Chapter 17*) have also been implicated. Clinical features of acute pancreatitis include attacks of severe abdominal pain that may extend to the back, vomiting, fever and shock. The leakage of pancreatic enzymes into the bloodstream from damaged tissue and a demonstration of increased plasma amylase activity aid in diagnosis. Typically, there is a fivefold increase in plasma amylase activity over the first two days of an attack, which then returns to normal levels within three to five days. Treatment is symptomatic and aimed at maintaining circulation and fluid volume and decreasing pain with analgesics.

Chronic pancreatitis is a slow, progressive destruction of pancreatic tissue accompanied by inflammation and fibrosis. Like acute pancreatitis, the damage is believed to be due to autodigestion of pancreatic tissue by the activation of enzymes *in situ*. Excessive alcohol intake is the leading cause of chronic pancreatitis, although many cases are idiopathic. Clinical features of chronic pancreatitis include severe and persistent abdominal pain, weight loss, malabsorption and hyperglycaemia. Complications of the condition include diabetes mellitus (*Chapter 7*). Chronic pancreatitis is investigated using X-ray examination of the pancreas to reveal calcification that may result from the release of free fatty acids following the breakdown of fats and by examining the faeces to detect steatorrhoea. Chronic pancreatitis results in pancreatic insufficiency, which may be assessed using the faecal elastase test. Pancreatic elastase can catalyse the hydrolyses of connective tissue proteins and its activity in faeces can be determined. Patients with more than 200 μg elastase/g of faeces have normal pancreatic function, whereas values less than this are associated with pancreatic insufficiency. Treatment of chronic pancreatitis is directed towards management of pain and rectifying the nutritional disorders that arise from malabsorption (*Section 11.10*).

Several tests are available to assess pancreatic function. One of which is the fluorescein dilaurate test, which indirectly assesses the activity of pancreatic enzymes. The patient is given an oral dose of the synthetic ester fluorescein dilaurate, which is hydrolysed to release fluorescein by pancreatic cholesterol esterase. Fluorescein, but not fluorescein dilaurate, is absorbed by the small intestine and transported to the liver where it is converted to fluorescein glucuronide, which is then excreted in the urine. The latter can be detected by its characteristic fluorescence. The test is controlled for variations in intestinal absorption, hepatic conjugation and renal excretion by repeating the test the next day but using an oral dose of fluorescein. The ratio of fluorescein excreted after administration of fluorescein dilaurate to that excreted after using free fluorescein is greater than 0.3 in normal individuals. However, a ratio less than 0.2 is indicative of abnormal pancreatic function. Ratios of 0.2–0.3 are inconclusive. Other investigations of pancreatic function include the *p*-aminobenzoic acid (PABA) test (*Figure 11.51*). The patient is given 0.5 g of the synthetic peptide benzoyltyrosylaminobenzoic acid

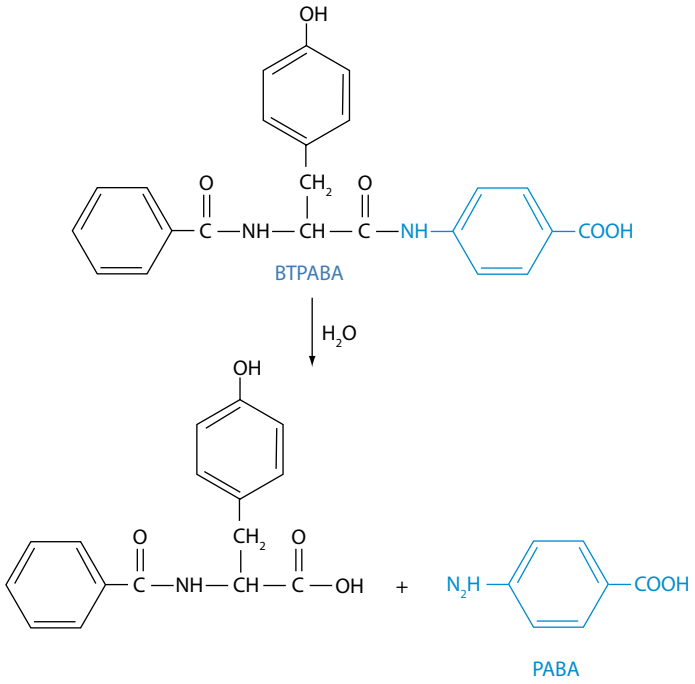


Figure 11.51 The hydrolysis of the synthetic peptide benzoyltyrosylaminobenzoic acid (BTPABA) by chymotrypsin activity to release *p*-aminobenzoic acid (PABA). ⏎

(BTPABA), which is hydrolysed to PABA in the small intestine by pancreatic chymotrypsin. Following its absorption, the PABA is excreted unchanged in the urine. The patient is also given a known amount of ^{14}C-PABA to correct for variations in its absorption, metabolism and excretion. The amount of PABA found in the urine is a measure of the activity of pancreatic chymotrypsin. There is reduced excretion of urinary PABA in individuals with abnormal pancreatic function.

Disorders of the liver, gallbladder and bile duct

Jaundice is the yellow discolouration of tissues due to an accumulation of bilirubin (see *Figure 11.5*). Many disorders of the liver give rise to jaundice, although clinical jaundice may not be seen until the concentration of bilirubin in the serum is greater than 50 μmol/L. The causes of jaundice can be prehepatic, hepatic or posthepatic.

Prehepatic jaundice can be caused by haemolysis, where an increased breakdown of haemoglobin produces such large amounts of bilirubin that it overloads the conjugating mechanism (*Chapter 13*). Such bilirubin is mostly unconjugated and commonly occurs in newborn babies. If the concentration of serum bilirubin approaches 200 μmol/L (normally values are less than 17 μmol/L for total and below 5.2 μmol/L for direct bilirubin), then phototherapy (*Chapter 6* and *Margin Note 11.7*) is used to degrade it; otherwise, its high concentration may lead to **kernicterus** and damage the brain.

Other causes of prehepatic hyperbilirubinaemia include haemolytic disease of the newborn due to Rhesus incompatibility (*Chapter 6*) and ineffective erythropoiesis, which occurs in pernicious anaemia (*Chapter 13*). The most common causes of hepatic hyperbilirubinaemia are viral hepatitis and paracetamol (acetaminophen) poisoning (*Chapter 12*). There is also physiological jaundice of the newborn, a mild unconjugated hyperbilirubinaemia that develops because of low activity of UDP glucuronyltransferase following birth. Activity increases within two weeks and the jaundice disappears. Other causes include Gilbert and Crigler–Najjar syndromes.

In Gilbert syndrome, the affected individuals have an inherited partial deficiency of hepatic UDP glucuronyltransferase. Patients present with a mild jaundice and occasionally suffer from abdominal discomfort but otherwise the condition is harmless. Fasting, infection, stress and excessive alcohol intake may aggravate the symptoms. Treatment of Gilbert syndrome is by administration of phenobarbitone to stimulate glucuronyltransferase activity. Crigler–Najjar syndrome is a rare hereditary disorder characterised by an absence of UDP glucuronyltransferase activity from birth. Patients suffer from severe unconjugated hyperbilirubinaemia. Treatment using phototherapy in affected newborns may temporarily reduce the unconjugated hyperbilirubinaemia, but infants generally

Margin Note 11.7 Phototherapy for jaundice in neonates ⏎

Phototherapy is the most common method of treating jaundice in neonatal babies. It involves exposing the baby to blue light (wavelength 400–500 nm). This is absorbed by the bilirubin, degrading it to a form that is more water soluble and can be excreted in the faeces and urine. A baby with jaundice may need repeated phototherapy over several days.

die within one year of birth. One cause of posthepatic hyperbilirubinaemia is **cholestasis**, that is, a failure of bile to reach the small intestine. Cholesterol is virtually insoluble in water and is maintained in an aqueous environment in vesicles combined with phospholipids and bile salts. In normal conditions, the vesicles maintain the concentration of cholesterol in bile near its saturation point. Cholesterol monohydrate crystals form when the ratio of cholesterol, phospholipids and bile salts exceeds the normal range and results in the formation of gallstones by **cholelithiasis**. Cholesterol forms approximately 80% of gallstones; the remaining 20% consists of Ca^{2+} and bilirubin. Gallstones vary in size from that of a grain of sand to the diameter of a golf ball. In many cases, the smaller stones can be excreted in the bile duct without causing harm. Larger gallstones usually cause abdominal pain and may be so large as to obstruct the flow of bile into the small intestine. However, in some cases, gallstones may exist for years without causing any symptoms. When there is a complete blockage, there is little or no urobilinogen in the faeces, which are pale coloured due to absence of stercobilinogens. When the blockage is removed, urobilinogen becomes detectable in the urine and the faeces regain their normal colour. Occasionally, intrahepatic obstruction arises, where a blockage affects the bile canaliculi in liver cirrhosis (*Chapter 12*) or cancer (*Chapter 17*). This type of blockage causes an increase in the concentration of conjugated bilirubin in the serum.

It is essential to determine whether the cause of the increased amounts of conjugated bilirubin is intra- or extrahepatic because it is of diagnostic significance and determines subsequent treatment. The degree of obstruction to the flow of bile is usually greater in extrahepatic cholestasis. Extrahepatic cholestasis may benefit from surgery to remove the gallbladder or to remove the gallstone (*Figure 11.52*). Nonsurgical treatments are preferred because of the inherent risks associated with surgery. Oral dissolution therapy with ursodiol and chenodiol (*Figure 11.53*), which are derived from bile salts, is effective in treating small, predominantly cholesterol gallstones. Treatment may be required for months to years before the gallstones are dissolved, but is preferred in patients who cannot undergo surgery.

Figure 11.52 Gallstones in an excised gallbladder. Courtesy of Dr H.W. Ray, Southwest Tennessee Community College, USA.

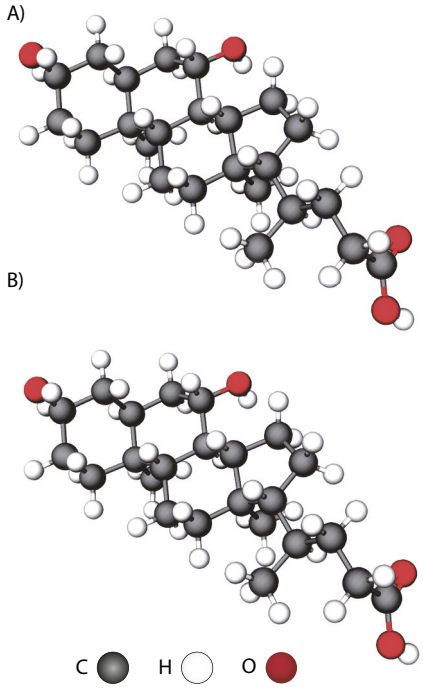


Figure 11.53 Computer drawn structures of (A) ursodiol and (B) chenodiol.

Hepatitis

Acute hepatitis is caused by infection and subsequent inflammation of the liver, where liver cells are destroyed and the liver becomes necrotic. The most common cause is viral infections, for example with hepatitis A, B, C, D and E viruses (HAV, HBV, etc.), although drugs, toxins and autoimmune reactions can also lead to acute hepatitis. The initial symptoms of acute viral hepatitis include malaise, anorexia, fever, rashes, abdominal pain, dark urine and jaundice. A clinical history of recent blood transfusions or intravenous drug use may all suggest acute hepatitis. Blood tests based on antigen–antibody reactions (*Chapter 4*) are conducted to establish the type of virus causing the hepatitis.

BOX 11.9 CHOLELITHIASIS AND CHOLESTASIS IN PREGNANCY

Cholelithiasis, or the presence of gallstones, in the gallbladder (see *Figure 11.52*) has a general incidence 1 in 1000, but pregnancy induces changes in the composition of bile that increases its frequency to 6 in 100. For example, in the second trimester, the bile salt pool generally decreases but biliary cholesterol concentrations may increase, producing bile that is more prone to form stones. Additionally, emptying of the gallbladder slows in the second trimester, further increasing the risk of cholelithiasis.

The symptoms of cholelithiasis are similar in pregnant and nonpregnant patients and may present as pain in the middle of the upper abdomen (which can become more severe on eating fatty foods), jaundice and fever. Symptoms usually occur only when the stones block one of the ducts in the biliary system. Cholelithiasis is frequently asymptomatic, and often the stones are only discovered by routine X-ray examination, by surgery or postmortem (autopsy). An ultrasound examination of the liver is helpful in determining the presence of gallstones. Surgical treatment can be safely accomplished in the first or second trimester but should be avoided during the third trimester because of the enlargement of the uterus.

Obstetric cholestasis (OC) or cholestasis of pregnancy is a liver disorder where the flow of bile from the liver in pregnant women is reduced. The condition affects 0.7% of pregnancies in multiethnic populations, but higher incidences of 1.5% and 4% have been reported in Scandinavia and Chile, respectively. Typically, OC presents in the third trimester, but can occur as early as the 13th week. In 80% of patients, it presents with severe pruritus (itching), with jaundice developing in the remaining 20%, showing typical dark urine and light-coloured stools. In a normal pregnancy, itching is not uncommon and is thought to be related to hormonal changes and stretching of the skin over the abdomen as the baby grows. However, in OC, the itching generally begins elsewhere, especially on the arms, legs, hands and soles of the feet, face, back and breasts. It is usually worse at night, leading to sleeplessness and exhaustion, and can be of such intensity that scratching draws blood. The itching completely disappears within a week or two of the birth and does not cause long-term health problems for the mother, but the condition is associated with increased health risks to the fetus. For example, studies in the USA have shown that intrahepatic OC is associated with a 12–44% incidence of prematurity and a 16–25% incidence of fetal distress, and increases the prenatal mortality rate from 1.3% to 3.5%. The pathophysiology of intrahepatic cholestasis of pregnancy remains poorly understood and, unfortunately, no one knows why babies are at risk of being stillborn. One possibility is that the liver cannot cope with the increased amounts of hormones produced during pregnancy, which reduces the flow of bile leading to a build up of bile salts in the blood.

In addition to a complete medical history and physical examination, generalised severe itching without a rash is often the first clue in diagnosing OC. This can be confirmed by liver function tests (LFTs; *Box 11.10*) and serum bile acid tests; the latter are the most sensitive. Normally, the amount of bile acids in blood increases before the LFTs can detect any changes. Blood tests to check blood clotting in OC are necessary prior to birth, and patients may require extra vitamin K, since a lack of the vitamin may decrease the effectiveness of clotting and increase blood loss during the birth (*Chapter 13*). Patients with OC require cardiotocography (that is, monitoring the heartbeat of the fetus over a set period of time), ultrasound scans and blood tests. Some pregnant women may be hospitalised to evaluate the progress of the fetus. Close fetal surveillance at delivery is also desirable.

Following a diagnosis of OC, patient care involves giving general support. Specific treatments are determined by the medical history, overall health and tolerance of the patient to specific medications and by the severity of the condition. Resting as much as possible and eating a well-balanced diet that includes large amounts of vegetables, fruit and wholewheat cereals, including bread, may help, as do frequent cold baths. Calamine lotion and loose cotton clothing are used to relieve the itching. Steroids may be used to reduce the levels of bile salts in the blood and relieve the itching. For example, treatment with ursodeoxycholate helps increase the flow of bile, reduce the level of bile acids in the blood and ameliorate the pruritus and is well tolerated by both mother and fetus. Dexamethasone (see *Figure 7.56*) is sometimes prescribed to increase the maturity of the fetal lungs before delivery and may also help relieve maternal itching. Parenteral vitamin K supplementation is recommended for patients with prolonged cholestasis and when blood clotting factor concentrations are abnormal.

Obstetric cholestasis may also increase the risk of postpartum maternal haemorrhage. If the wellbeing of the mother or the fetus is judged to be at risk, then an early delivery at week 37 or 38 may be necessary. There does not appear to be any harmful effects to babies born to mothers with OC. Maternal symptoms usually resolve within two days of delivery and serum bilirubin and LFTs soon return to normal values after delivery. Obstetric cholestasis is not thought to cause any lasting liver damage, although it may leave the liver more sensitive to normal changes in the concentration of hormones, leading to bouts of mild itching during the menstrual cycle just before ovulation or just prior to the start of a period. A consultant obstetrician familiar with the condition should carefully manage any subsequent pregnancies, since the condition recurs in 60–70% of cases, but it may not follow the same pattern. For example, the itching may be more severe and could begin earlier in the pregnancy.

Many patients present with proteinuria and bilirubinuria and show increased serum alkaline phosphatase activity. A liver biopsy can confirm an initial diagnosis.

Hepatitis A virus (*Figure 11.54*) causes a mild hepatitis from which patients recover usually with no complications. The virus is transmitted by contaminated food or drink, especially where sanitation is poor. Following an incubation period of 15–40 days, the patient develops fever, sickness and, shortly afterwards, jaundice. Its treatment includes bed rest and is aimed at alleviating the symptoms, but intramuscular injections of γ globulins have been used.

Hepatitis B virus infections are more serious than HAV with a mortality rate of 5–20%, although most patients gradually recover. Hepatitis B virus spreads from one person to another via body fluids, such as blood, saliva, semen, vaginal fluids, tears, breast milk

and urine. Transmission may occur during sexual activity with an infected person and vertically from an infected mother to the baby. It is commonly present in drug addicts. The symptoms develop suddenly after an incubation period of one to six months and include fever, chills, weakness and jaundice. Its therapy is similar to that of HAV.

In contrast with other types of hepatitis, more than 80% of HCV infections cause chronic liver disease. In 2023, 40 million people worldwide were infected with HCV. Blood transfusions were the most common means of transmission prior to the testing of blood products for HCV. The infection is mild in the early stages and is often only diagnosed when it has already caused severe liver damage. For this reason, infection with HCV has been referred to as the 'silent epidemic'. Treatment of HCV, like HAV and HBV, is bed rest with adequate fluid and food intakes and drug therapy using interferon-α (*Chapter 4*). However, recent years have seen the development of a battery of low M_r antiviral drugs, such as sofosbuvir, elbasvir, glecaprevir and velpatasvir, which interfere with viral replication at different stages in its replication cycle (*Chapter 2*). All are used in a variety of combinations to increase effectiveness and minimise resistance developing.

The HDV occurs only with or after infection with HBV and its mode of transmission is identical to that of HBV. Hepatitis E was initially grouped as a type C virus. It occurs in people who have been to parts of the world where this virus is endemic, such as India. It is transmitted by water contaminated with faecal material.

Infections with HBV and HCV are associated with liver cancer (*Chapter 17*). A further serious complication of many cases of acute hepatitis is the development of chronic hepatitis, an inflammation of the liver that persists for more than six months without improvement. Its causes include autoimmune liver damage, chronic infection with HBV and excessive drug and alcohol use. Chronic hepatitis can be divided into two histological types, namely chronic persistent hepatitis, which has a good prognosis, and chronic active hepatitis that may respond to immunosuppressive or antiviral agents but often progresses to cirrhosis, leading to death within five years as a result of liver failure.

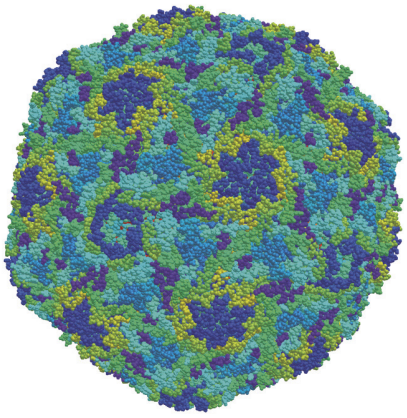


Figure 11.54 Molecular model of a hepatitis A virus capsid. PDB file 5WTE. ⏎

Cirrhosis

Cirrhosis (*Chapter 12*) is a condition where the liver responds to injury or death of some of its cells by producing strands of fibrous tissue between which are nodules of regenerating cells (see *Figure 12.19*). Patients with cirrhosis may be asymptomatic for a long period of time before vague symptoms such as nausea, vomiting, anorexia, weakness, weight loss and oedema of the legs become apparent. Its clinical complications (*Figure 11.55*) include jaundice, **ascites** (an abnormal accumulation of fluid in the abdomen), GIT bleeding and hepatic encephalopathy. Associated endocrine dysfunction (*Chapter 7*)

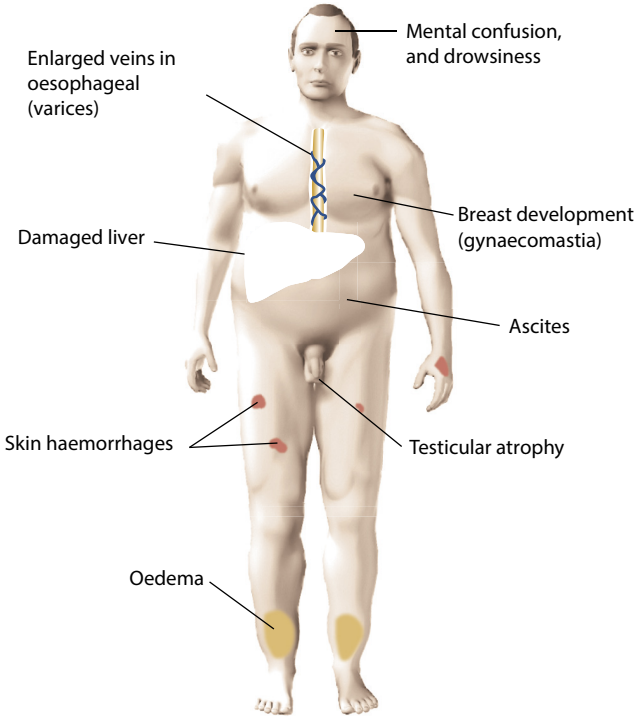


Figure 11.55 Clinical problems associated with liver cirrhosis. ⏎

can lead to testicular atrophy, gynaecomastia and hypothermia. Cirrhosis may interfere with intrahepatic circulation causing gradual failure of liver function.

Cirrhosis can be divided into three types:

1. alcoholic
2. postnecrotic
3. biliary

Alcoholic cirrhosis is discussed in *Chapter 12*. Postnecrotic cirrhosis accounts for about 25% of all cases of cirrhosis and is associated with viral infections, the use of certain drugs and poisons. About 25% of postnecrotic cirrhosis cases have a prior history of viral hepatitis. Unfortunately, 75% of all patients with postnecrotic cirrhosis die within one to five years. Biliary cirrhosis accounts for approximately 15% of all cases of cirrhosis and is characterised by the death of liver cells surrounding bile ducts. It is mostly caused by an obstruction of the bile duct leading to an accumulation of bile within the liver.

Diagnosing cirrhosis involves abdominal palpation and X-raying the abdomen, which often reveal an enlarged liver. A liver biopsy is required to confirm the diagnosis. Other laboratory tests may reveal anaemia or hyperbilirubinaemia, and LFTs investigate any increases in the activities of several enzymes (*Box 11.10*). There are no drugs that can arrest or reverse the fibrotic process in cirrhosis; treatment is aimed at dealing with the underlying cause(s), for example alcohol abuse or biliary obstruction, and by treating any complications. In extreme and suitable cases, a liver transplant can be attempted (*Chapter 6*). These are often successful but, when the initial clinical condition arose from substance abuse, patients must not relapse otherwise the new liver will follow the fate of the old.

BOX 11.10 LIVER FUNCTION TESTS

The concentration of serum albumin in plasma, a number of its enzyme activities and the prothrombin time can all be used to assess liver function.

Albumin is a major plasma protein synthesised in the liver, and its concentration in blood reflects the functional capacity of the liver. It is low in chronic liver disease but tends to be normal in the early stages of acute hepatitis before decreasing.

Damage to liver cells releases a number of enzymes, altering their activities in the plasma. The enzymes monitored in LFTs include:

- alanine and aspartate transaminases (ALT and AST, respectively)
- alkaline phosphatase (ALP)
- γ-glutamyltranspeptidase (GGT)

Aspartate transaminase occurs in both the cytosol and the mitochondria of hepatocytes, whereas ALT is found only in the cytosol. Alanine transaminase is restricted to the liver, whereas AST is also found in pancreatic, skeletal and cardiac muscle tissues. In hepatocellular damage, activities of AST and ALT in the blood may increase 10-fold but, in obstructions of the bile duct or cholestasis, the increases may be relatively slight; usually no more than two to three times their normal levels. Alanine transaminase and AST measurements are useful in monitoring the progress of hepatocellular damage when decreasing activities in blood samples suggest an improvement in the disease. Alkaline phosphatase is found on the surface of hepatocytes and in the microvilli of bile ducts, but is not specific to the liver. Its activity is increased in cholestasis. In hepatocellular disease, ALP activities may be normal or slightly increased. A decrease in plasma ALP activity suggests a correction of cholestasis and may be useful for monitoring this condition. Plasma GGT activity increases in both hepatocellular disease and cholestasis. The assay for this enzyme is sensitive but not specific for liver disease, as activity is increased with some drug therapies and by alcohol intake.

The prothrombin time is a test of the clotting activity of the plasma, which reflects the status of vitamin K-dependent clotting factors (*Chapter 10*). An increase in prothrombin time is an early feature of acute liver disease (or a vitamin K deficiency).

11.12 CASE STUDIES

CASE STUDY 11.1

Alice, a five-year-old child, presented with loss of weight, abdominal distension and obvious signs of anaemia. Her stools were loose, bulky and pale coloured and had an offensive smell. A variety of clinical tests were ordered with the following results (reference ranges are shown in parentheses):

Serum albumin	29	(32–48 g/L)
Serum iron	4	(10–30 µmol/L)
Xylose absorption	0.5	(>1.2 g/5 h)
Faecal fat	29	(<15 g/3 days)

Jejunal biopsy showed villous atrophy.

Questions

(a) What is the most likely diagnosis?
(b) How should Alice be treated?

CASE STUDY 11.2

It was noticed that Sadaf, a fine art student recovering from an attack of influenza, was slightly jaundiced. Worried that she may have hepatitis, her doctor asked her to provide some blood and urine for laboratory investigations. Test results are shown below (reference ranges are shown in parentheses):

Serum bilirubin	60	(3–20 μmol/L)
Serum ALP	74	(30–90 IU/L)
Serum AST	35	(10–50 IU/L)
Haemoglobin	140	(115–155 g/L)
Reticulocyte count	1.5	(1–2%)
Urine bilirubin	Negative	(negative)

Questions

(a) What is the most probable diagnosis to account for these signs and symptoms?
(b) Suggest how Sadaf's condition should be managed.

CASE STUDY 11.3

Mark, a 52 year old plumber, was admitted to hospital because of severe abdominal pain. The pain had started suddenly about 15 hours previously. On questioning, Mark admitted to being a heavy alcohol drinker over many years. Clinical examination showed that Mark had a tender abdomen and was in mild shock. Radiographic examination did not show any evidence of intestinal obstruction. Biochemical tests gave the following results (reference ranges are shown in parentheses):

Serum Na^+	140	(135–145 mmol/L)
K^+	4. 2	(3.6–5.0 mmol/L)
Urea	11	(3.3–6.7 mmol/L)
Glucose	13	(2.8–6.0 mmol/L)
Amylase	5000	(<300 IU/L)

Questions

(a) What is the most likely diagnosis?
(b) What is the underlying causes of Mark's high serum urea concentration?
(c) How should Mark's clinical problems be treated?

CASE STUDY 11.4

XiaoFang, a 32 year old from China, entered England to study as a mature research student. Wishing to 'integrate' fully in her new environment, she adopted many of the dietary habits typical of the country. After six healthy months, she unfortunately suffered a bout of viral gastroenteritis that required hospitalisation. XiaoFang, fortunately, made a complete recovery and was able to continue her new lifestyle. However, she again became ill, but with severe indigestion, abdominal cramps, intestinal bloating and periods of diarrhoea. Symptoms were most pronounced about one hour after breakfast.

Questions

(a) Suggest the most probable cause of XiaoFang's symptoms.
(b) What test(s) could be used to help your initial diagnosis? Include any precautions that may be necessary in your clinical investigations.
(c) Assuming your initial diagnosis is correct, what treatment do you recommend for XiaoFang?

11.13 SUMMARY

The GIT and its associated organs are necessary for the digestion of food and the absorption of nutrients derived from food. Digestion takes place in the mouth, oesophagus, stomach and small intestine, with the bulk of ingested water being absorbed in the latter. The pancreas provides enzymes for digestion in the small intestine, while the liver secretes bile, containing detergents that aid in the digestion of fats. Most nutrients from

digestion are carried to the liver in the hepatic portal vein. The GIT also contributes to nutrient homeostasis, through metabolic reactions that take place in GIT enterocytes and the delivery of their metabolic products, such as lactate, to the liver. The GIT also produces hormones that help in coordinating its activities. Disorders of the GIT include gastritis and ulcers that are associated with infection of *H. pylori*, intolerance to certain nutrients such as lactose, coeliac disease and malabsorption. Clinical problems associated with the pancreas, such as acute and chronic pancreatitis, are serious conditions in that they may be fatal. Disorders of the liver may be caused by drugs, poisons, viruses, alcohol or blockage of the bile duct, as in cholestasis. Liver disorders may be diagnosed by abdominal palpation to detect a swollen liver, X-ray examination and LFTs. Jaundice is a frequent complication of liver disease.

QUESTIONS

1. Which of the following is the odd ONE out?
- **(a)** Cementum.
- **(b)** Crown.
- **(c)** Dentine.
- **(d)** Gingivitis.
- **(e)** Periodontal ligament.

2. Which of the following statements is/are true?
- **(a)** Digestion involves the hydrolysis of polysaccharides, oligosaccharides, proteins, and fatty and nucleic acids.
- **(b)** The mesentery and mesocolon are folds of the peritoneum.
- **(c)** The pancreas functions as both an exocrine and an endocrine gland and in the absorption of lipids.
- **(d)** The hepatic portal vein and hepatic artery supply liver sinusoids with blood, which drains directly into the inferior vena cava.
- **(e)** The bile salts sodium glycocholate and taurocholate are detergents produced in the gallbladder that aid in the digestion of lipids by emulsifying them to form water-soluble complexes.

3. Which of the following statements is/are true?
- **(a)** Digestive lipases are produced in the mouth, stomach and pancreas.
- **(b)** Liver sinusoids are lined with Kupffer cells that form part of the mononuclear phagocytic system (reticuloendothelial system).
- **(c)** UDP glucuronyltransferase catalyses the conjugation of a variety of molecules with glucuronic acid.
- **(d)** The submandibular glands are sited under the jaws and the sublingual and parotid salivary glands under the tongue.
- **(e)** Pancreatic lipase hydrolyses lipids to fatty acids and 1- and 3-monoacylglycerols but requires the presence of colipase to support these activities.

4. Which ONE of the following statements is true?
- **(a)** Parietal or oxyntic cells secrete pepsinogen.
- **(b)** Peptic or chief cells produce HCl.
- **(c)** Gastric lipase has a pH optimum of 5–7.
- **(d)** The cardiac sphincter controls the release of chyme into the duodenum.
- **(e)** Exopeptidases and endopeptidases catalyse the release of free amino acids from peptides and proteins.

5. Which of the following is the odd ONE out?
- **(a)** Pepsin.
- **(b)** Na^+/K^+-ATPase.
- **(c)** Sucrase.
- **(d)** Trypsin.
- **(e)** Pancreatic lipase.

6. Which of the following is associated with the development of peptic ulcers?
- **(a)** Smoking.
- **(b)** *Helicobacter pylori*.
- **(c)** Weight loss.
- **(d)** Coeliac disease.
- **(e)** Lactose intolerance.

7. Which ONE of the following tests is best used to assess pancreatic function?
   (a) Faecal fat test.
   (b) Serum bilirubin determination.
   (c) Oral glucose tolerance test.
   (d) Fluorescein dilaurate test.
   (e) Xylose absorption test.
8. List some possible causes of jaundice.
9. List four causes of malabsorption.

FURTHER READING

Adams SM *et al*. (2022) Ulcerative colitis: rapid evidence review. *Am Fam Physician* **105**:406–411.

Ammouri W *et al*. (2020) Pernicious anaemia: mechanisms, diagnosis, and management. *EMJ Hematol (US ed)* **1**:71–80.

Catassi C *et al*. (2022) Coeliac disease. *Lancet* **399**:2413–2426.

Clark R and Johnson R (2018) Malabsorption syndromes. *Nurs Clin North Am* **53**:361–374.

De Angelis C *et al*. (2018) Diagnosis and management of Zollinger-Ellison syndrome in 2018. *Minerva Endocrinol* **43**:212–220.

Devarbhavi H *et al*. (2023) Global burden of liver disease: 2023 update. *J Hepatol* **79**:516–537.

Ford AC *et al*. (2017) Irritable bowel syndrome. *N Eng J Med* **376**:2566–2578.

Ginès P *et al*. (2021) Liver cirrhosis. *Lancet* **398**:1359–1376.

Guéant J-L *et al*. (2022) Vitamin B12 absorption and malabsorption. *Vitam Horm* **119**:241–274.

Hemmer A *et al*. (2023) Inflammatory bowel disease: a concise review. *S D Med* **76**:416–423.

Jeppesen PB (2014) New approaches to the treatments of short bowel syndrome-associated intestinal failure. *Curr Opin Gastroenterol* **30**:182–188.

Johnson LR (2018) *Gastrointestinal Physiology*, 9th edition. Mosby, St Louis, USA.

Kiela PR and Ghishan FK (2016) Physiology of intestinal absorption and secretion. *Best Pract Res Clin Gastroenterol* **30**:145–159. doi: 10.1016/j.bpg.2016.02.007.

Lanas A and Chan FKL (2017) Peptic ulcer disease. *Lancet* **390**:613–624.

Littlefield A and Lenahan C (2019) Cholelithiasis: presentation and management. *J Midwifery Womens Health* **64**:289–297.

Louw JA and Marks INS (2004) Peptic ulcer disease. *Curr Opin Gastroenterol* **20**:533–537.

Mederos MA *et al*. (2021) Acute pancreatitis: a review. *JAMA* **325**:382–390.

Milman NT (2020) A review of nutrients and compounds, which promote or inhibit intestinal iron absorption: making a platform for dietary measures that can reduce iron uptake in patients with genetic haemochromatosis. *J Nutr Metab* **2020**:7373498. doi: 10.1155/2020/7373498.

Ormond M *et al*. (2015) Disorders of the mouth. *Medicine* **43**:187–191.

Pavlovic Markovic A *et al*. (2022) Jaundice as a diagnostic and therapeutic problem: a general practitioner's approach. *Dig Dis* **40**:362–369.

Pelaseyed T and Hansson G (2020) Membrane mucins of the intestine at a glance. *J Cell Sci* **133**:jcs240929. doi: 10.1242/jcs.240929.

Sebastián Domingo JJ (2022) Irritable bowel syndrome. *Med Clin (Barc)* **158**:76–81.

Seth D *et al*. (2020) Food allergy: a review. *Pediatr Ann* **49**:e50–e58.

Seyedian SS *et al*. (2019) A review of the diagnosis, prevention, and treatment methods of inflammatory bowel disease. *J Med Life* **12**:113–122.

Tack J *et al*. (2004) Pathophysiology and treatment of functional dyspepsia. *Gastroenterology* **127**:1239–1255.

Toca MDC *et al*. (2022) Lactose intolerance: myths and facts. An update. *Arch Argent Pediatr* **120**:59–66.

Torres J *et al*. (2017) Crohn's disease. *Lancet* **389**:1741–1755.

Vege SS and Chari ST (2022) Chronic pancreatitis. *N Engl J Med* **386**:869–878.

Vila T *et al*. (2019) The power of saliva: antimicrobial and beyond. *PLOS Pathog* **15**:e1008058. doi: org/10.1371/journal.ppat.1008058.

Waldum H and Fossmark R (2021) Gastritis, gastric polyps and gastric cancer. *Int J Mol Sci* **22**:6548.

Toxicology

12

OBJECTIVES

After studying this chapter you should be able to:

- define the terms drug, xenobiotic, poison, toxin and toxicology
- outline the absorption, distribution and excretion of drugs
- explain the roles of the liver and kidneys in detoxification
- outline the types and clinical effects of some common poisons
- describe the general methods for treating and managing drug overdoses and poisoning

12.1 INTRODUCTION

The human body must be prepared for a daily onslaught of thousands of chemicals. Some will be food materials (*Chapter 10*), which are absorbed and metabolised because cells have membrane channels and enzymes to recognise and deal with these compounds. Other materials may or may not enter cells but, if they do, the body can recognise such compounds and must deal with them appropriately to prevent deleterious effects. Detoxification is not always successful in this aim, and many xenobiotics cause diseases and may be fatal if a threshold dose is exceeded. **Toxicology** is the study of the adverse effects of chemicals on organisms, that is, the study of poisons, where a poison can be defined as any substance that causes injury, illness or death. Toxicology covers the study of the adverse effects of chemicals including drugs, chemicals acquired from the environment and toxins, which are defined here as harmful substances produced by other organisms, often derived from microorganisms, for example the bacterial toxins described in *Chapter 2*.

Xenobiotics, literally meaning 'stranger to life', are substances that do not originate in the body but are pharmacologically, endocrinologically or toxicologically active. Thus, they might be drugs or synthetic chemicals, or a substance produced in one organism and introduced into another where it would not normally occur. An example would be a compound produced by a plant and ingested as food. Equally, it might be a material that has been synthesised chemically and has harmful effects, for example a poison or a carcinogen, or it could be beneficial, such as a medicinal drug. Drugs are xenobiotics that are used to achieve certain effects. For example, paracetamol alleviates headaches and aspirin mediates inflammatory responses. Alternatively, a drug may also be a compound with pharmacological activity used for 'recreational' purposes or taken by an addict. The body must deal with these drugs and eventually eliminate them. An overdose may exceed the body's capacity to detoxify these compounds, with potentially disastrous effects. However, if the body inactivates a therapeutic drug too quickly, its effects will be short-lived. These are all considerations that pharmaceutical companies need to address when developing a new drug.

12.2 DRUG ACTION, DISTRIBUTION AND EXCRETION

Drugs are xenobiotics and may be defined as any substance, other than food, that affects a living process. **Pharmacology** is the study of the effects of drugs in the prevention, diagnosis and treatment or cure of disease. Such drugs are often referred to as

DOI: 10.1201/9781003296836-12

BOX 12.1 2,4 DINITROPHENOL, A RECURRING XENOBIOTIC

Despite its relative simplicity and low M_r, the xenobiotic 2,4-dinitrophenol (DNP; *Figure 12.1*) is a potent uncoupler of oxidative phosphorylation, that is, it breaks the link between the oxidation of reduced coenzymes by the electron transport chain and the production of ATP at the inner mitochondrial membrane. Ingestion of DNP is poisonous and can induce hyperthermia, tachycardia (*Chapter 14*), tachypnoea and metabolic acidosis (*Chapter 9*), effects that can be fatal.

During World War I, it was used as a precursor for the explosive trinitrotoluene (TNT) in the armaments industry. However, many workers exposed to DNP lost weight, suffered fatigue, sweated excessively and had increased body temperatures. In a number of cases, deaths occurred before appropriate safety measures were introduced. Given that it severs the link between the oxidation of fuel molecules, carbohydrates and triacylglycerols, and ATP formation, the use of DNP as a drug to treat obesity was explored in the early 1930s but without success due to its toxicity and the deaths of some recipients. A lethal oral dose may be as low as 4.3 mg/kg, while serum concentrations associated with fatalities are normally greater than 25 mg/L.

Unfortunately, in recent decades, it has become available illegally online for its illicit use by some body builders, people wishing to rapidly lose weight and some with eating disorders (*Chapter 10*). This has resulted in cases of DNP poisoning and deaths due its side effects. Indeed, in 2019, Public Health England (now replaced by the UK Health Security Agency and the Office for Health Improvement and Disparities) issued an alert reference to NHS frontline staff on the need to remain vigilant for cases of DNP poisoning.

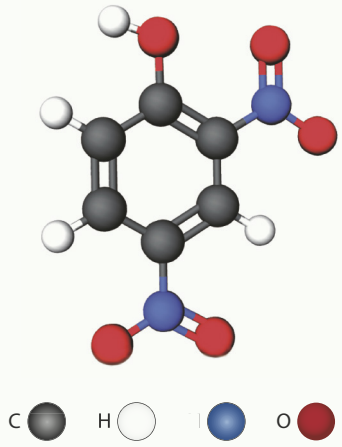
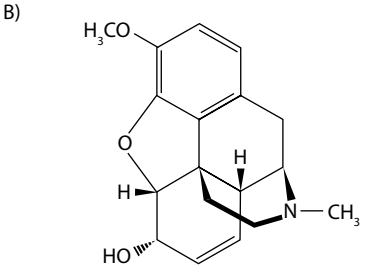
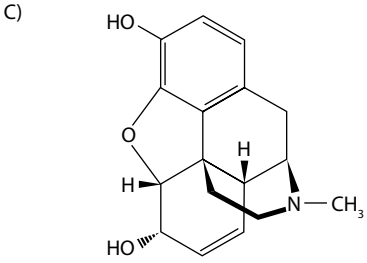
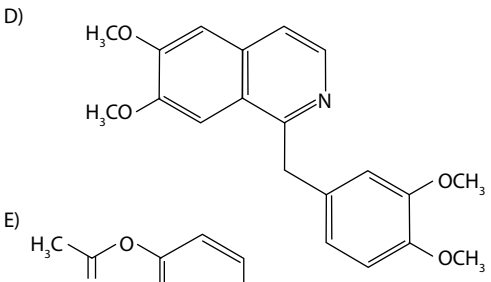


Figure 12.1 Computer drawn model of 2,4-dinitrophenol (DNP).







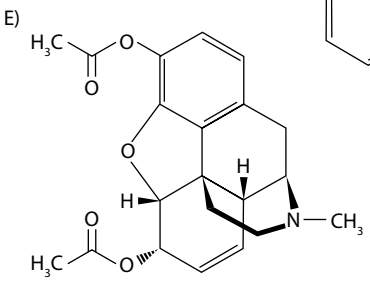


Figure 12.2 (A) The opium poppy, *Papaver somniferous*. (B) Codeine, (C) morphine, (D) papaverine and (E) heroin.

Margin Note 12.1 Opium

Opium is the milky-coloured exudate harvested from vertical cuts in the unripe fruits of the opium poppy (*Papaver somniferous*). It contains several bioactive compounds, including codeine, morphine and papaverine (*Figure 12.2*). Codeine is used clinically to treat mild–moderate pain and as an antitussive (suppresses coughing). Morphine is a powerful analgesic and is used therapeutically, mainly to give relief from intense pain. Its repeated use can lead to dependence. Papaverine is used to treat cardiovascular and abdominal spasms. Heroin, a diacetyl derivative of morphine, does not have any therapeutic value but is well known as a highly addictive substance of abuse. In the UK, it is classified as a class A drug and its use or sale is illegal; other counties have similar restrictions.

medicines, which distinguishes them from other drugs that are used for pleasure, such as some narcotics.

Pharmacotherapeutics is the branch of pharmacology concerned with the administration of drugs for prevention and treatment of disease. Drugs can be classified according to their chemical structure but, more often, in terms of their pharmacological effects. For example, they can be divided into three groups: chemotherapeutic drugs, for example the antibiotics described in *Chapter 3*, which are used to treat infectious diseases; pharmacodynamic drugs, such as sedatives, which are used in the treatment of noninfectious diseases; and a number of miscellaneous agents including narcotics and analgesics.

Any single drug may have a chemical, a brand and a generic name. The chemical name is given according to the rules of chemical nomenclature, whereas the brand name is supplied by the manufacturer. The generic name is a common, established name for a drug irrespective of that of its manufacturer.

Most drugs act on cells to alter a biological function. This pharmacological effect occurs as a consequence of the drug reacting with a receptor that controls a particular function, or because the drug alters a physiological mechanism that affects that function. The position at which the drug acts to produce a pharmacological effect is called its *site of action*. The mechanism of action of the drug is the biochemical or physiological process that occurs at this site to produce the desired effect. Some drugs have molecular structures resembling those of endogenous molecules and compete with them for binding sites, and act by preventing the formation, release, uptake or transport of key substances in the body or by forming complexes with biomolecules that can then activate receptors. Drug receptors are varied. They include enzymes, for example 3-hydroxy-3-methylglutaryl-coenzyme A reductase (HMG-CoA reductase), which is inhibited by drugs called statins, such as atorvastatin and simvastatin (*Figure 12.3*), used in the treatment of hypercholesterolaemia (*Chapter 14*). Others bind to nonenzymic proteins; for example, the H^+ transport pump of parietal cells in the stomach is inhibited by drugs like lansoprazole and esomeprazole used in treating gastric reflux (*Chapter 11*). The anthelmintics albendazole and mebendazole immobilise parasitic worms by binding to β tubulin monomers and so inhibit essential rearrangements of the parasite's cytoskeleton. Not all receptors are proteins: clofazimine, which is used to treat leprosy and tuberculosis, binds to DNA, preventing its replication and transcription.

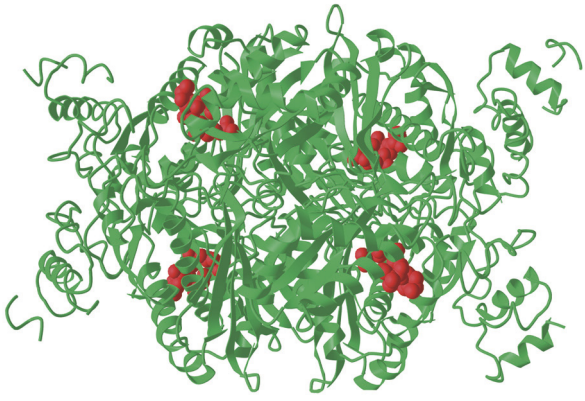


Figure 12.3 Molecular model of the catalytic portion of 3-hydroxy-3-methylglutaryl-coenzyme A reductase (HMG-CoA reductase), shown in green, with four bound simvastatin molecules, highlighted in red. PDB file 1HW9. ⏎

Drug–receptor interactions

The binding of a drug to its receptor usually depends on relatively weak effects, such as van der Waals forces and hydrogen and ionic bonds. Thus, the formation of the drug–receptor complex that elicits the response is normally freely reversible. Hence, the response to any drug is not usually permanent.

For many drugs, the extent and duration of the pharmacological effect are proportional to the concentration of the drug at the receptor. Indeed, a dose–response relationship exists between the concentration of drug in the serum and the pharmacological effect. This response eventually reaches a maximum effect because the receptor becomes saturated with the drug (*Figure 12.4*). The therapeutic range is the concentrations of drug

in the serum that are appropriate for clinical use. The dosage of any drug is planned to give a serum concentration within its therapeutic range. Therapeutic drug monitoring is often necessary to determine which given doses of a drug result in serum concentrations within the therapeutic range. The serum concentration of the drug must not fall below its minimum effective concentration (MEC) otherwise it will be ineffective. However, nor should it rise above its minimum toxic concentration (MTC) because of the danger of metabolic or structural damage.

The time required for the concentration of a drug in the blood to decline to half its original value is referred to as its **half-life** ($t_{1/2}$).

Pharmacodynamics and pharmacokinetics

It is essential that a number of properties relating to a medicinal drug, for example its **pharmacodynamics** and **pharmacokinetics**, are first determined prior to its therapeutic use. Pharmacodynamics describes how the drug interacts with its target site and the biochemical and physiological processes that result in any therapeutic or toxic effects. Pharmacokinetics relates to the uptake, distribution, metabolism and excretion from the body.

Administering medicines

For convenience, most medicines are given orally, although they may also be administered intravenously, intramuscularly or subcutaneously. When given orally, the absorption of the drug depends on its ability to disassociate from its dosing form, dissolve in gastrointestinal fluids and diffuse across the gastrointestinal tract (GIT) into the blood. The rate and extent of drug absorption varies with the nature of the drug, the matrix in which it is dissolved and the region of the GIT where it is absorbed. The proportion of the drug absorbed into the circulation is referred to as its bioavailability. For an orally delivered drug, this should generally be greater than 70% to be of therapeutic use. When the site of action is the GIT lumen itself, for example when treating a GIT infection, a low bioavailability would be advantageous.

A number of drugs undergo what is referred to as first-pass metabolism. They are absorbed rapidly and completely by the GIT but, nevertheless, have low bioavailability because they are transported to the liver in the hepatic portal vein and metabolised (*Section 12.3*) and therefore do not enter the systemic circulation. Drugs with delayed absorption are sometimes required, and special slow- or sustained-release formulations have been developed for these cases. Such drugs can be taken orally at less frequent intervals. Certain diseases that affect the GIT and the interaction of some drugs and foods in the GIT can delay their absorption.

Following absorption, drug distribution occurs when the compound enters the vascular system (*Chapter 14*). The physical, chemical and molecular properties of the drug can all influence its distribution. Its distribution may also be influenced by its binding to blood components and receptors and its ability to dissolve in lipids and pass through biological membranes. Many drugs bind to plasma proteins, and often an equilibrium is established between the protein-bound and the free drug. Only the free fraction can interact with receptors or cross cellular membranes. Any factor that changes drug–protein interactions may alter the distribution, pharmacological effects and excretion of the drug.

Drugs are excreted from the body by the biliary, GIT, pulmonary and/or renal routes. Most drugs are excreted through the renal system (*Chapter 8*) and, therefore, alterations in renal function may influence the half-life and serum concentration of the drug. A decline in renal function causes an increase in the concentration of the drug in serum, with an associated increased pharmacological effect.

12.3 DETOXIFICATION MECHANISMS

Xenobiotics, whether drugs or toxins, are normally absorbed through the lungs, GIT or skin. In addition, drugs may be injected through subcutaneous intramuscular, intraperitoneal or intravenous routes. Following absorption, xenobiotics are distributed around the body and a proportion will be absorbed by cells, while some may be directly excreted by the kidneys. Humans also have a number of physiological mechanisms devoted to detoxifying ingested xenobiotics, which largely involve enzyme activities in the liver and kidneys. Many of the biochemical detoxification mechanisms convert the xenobiotic to more water-soluble compounds that are more easily excreted. These mechanisms can sometimes prove inadequate. Also, in some unfortunate cases, the detoxification mechanisms form compounds that are even more toxic or carcinogenic than the original compound (see *Box 12.2*).

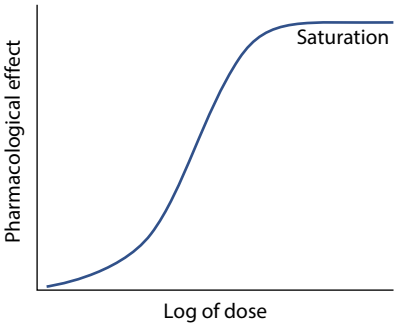


Figure 12.4 A dose–response curve for a typical drug. Saturation occurs when all of its receptors are occupied by the drug.

Margin Note 12.2
A lethal bacterial toxin

The bacterium *Clostridium botulinum* can cause a fatal type of food poisoning. It produces several toxins (see *Figure 2.27*) when growing on food in anaerobic conditions. In addition, its spores are resistant to heat; hence, food that has not been adequately sterilised and then stored in an anaerobic environment, such as bottled or canned foods, can give rise to food poisoning. These toxins inhibit the release of acetylcholine at neuromuscular junctions, leading to a flaccid paralysis in tissues served by susceptible neuromuscular and peripheral nerves (*Chapter 2*). The lethal dose of botulinum toxin for humans is not known with certainty, but estimates from studies with other primates suggest that approximately 0.1 µg given intravenously or intramuscularly and about 1µg if inhaled would kill a 70 kg human.

Surprisingly, given their extreme toxicity, toxins from *C. botulinum* are used cosmetically and therapeutically. Injections of 'Botox', the A toxin, into the skin of the face is used to relax the face muscles, reducing the wrinkled appearance of ageing (*Chapter 18*). The injections must be repeated every few months, as the effect is temporary. Botulinum toxin is also used in clinical practice to control conditions associated with inappropriate secretions of acetylcholine, for example the muscular spasms associated with the neurological disorder dystonia.

Treatment for poisoning involves administering an antidote to the toxin or drug when one is available. For many common poisons, this is not the case and therapy involves general supportive measures, such as decreasing their absorption from the GIT, increasing the rate of elimination from the body or altering the distribution within the body to protect susceptible tissues. An optimal therapeutic concentration of a medicinal drug can only be maintained in the plasma if the patient fully complies with the prescribed dose. Unfortunately, the most common reason for emergency admissions to hospitals is because of an excessive intake of a drug prescribed for medication. Other causes of poisoning may be accidental, suicidal or homicidal in intent.

Detoxification by the liver and kidneys

Many poisons are lipophilic and are only sparingly soluble in water and so cannot easily be excreted by the kidneys. Thus, a major part of detoxification in the liver is to convert them to compounds that have increased water solubility and are more readily removed by the kidneys. Detoxification reactions in the liver are favoured by its microstructure, which consists of lobules with a central vein and peripheral branches of the hepatic artery and the hepatic portal vein. Blood leaves branches of the hepatic artery and hepatic portal vein and percolates from the periphery of the lobule through sinusoids, bathing the hepatocytes of the lobule as it flows (*Figure 12.5*). While it is unusual for blood to come into such close contact with tissue cells, this arrangement allows poisons to be removed as the blood flows through the lobule to be collected by branches of the hepatic vein. Also, mitotic divisions within the organ replace those liver cells that are irreparably damaged. It goes without saying that a disease of the liver can compromise its ability to deal with toxic substances, which can lead to significant clinical consequences.

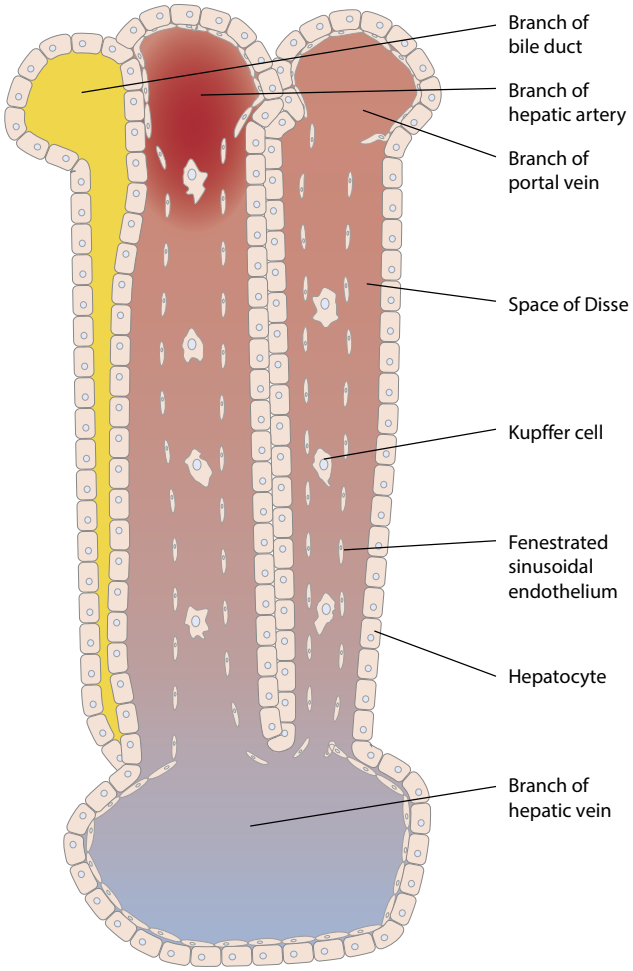


Figure 12.5 Schematic of a portion of a liver lobule. Blood enters at the periphery from a branch of the hepatic portal vein and the hepatic artery and eventually leaves at the centre of the lobule via a branch of the hepatic vein. Close contact of the blood with the liver cells means that these cells are the first to contact poisons and toxins after their absorption from the gastrointestinal tract (GIT) and transport in the portal system. The hepatocytes contain enzymes involved in detoxification.

Detoxification by hepatocytes

Detoxification is achieved in two phases: phases I and II. Both phase I and II enzymes have broad substrate specificities and are therefore able to detoxify a wide range of organic toxins and drugs. This is essential, since the range of xenobiotics to which the body may be exposed is enormous and specific enzyme systems could not be available to deal with each one individually.

In phase I, the xenobiotic is oxidised and/or hydroxylated by two mixed function oxidases (MFOs)/monooxygenase systems. One system is a flavin-containing monooxygenase (FMO) and the other a cytochrome P-450 monooxygenase. Both systems require NADPH as a coenzyme and use dioxygen, and are localised in the smooth endoplasmic reticulum (SER; *Figure 12.6*).

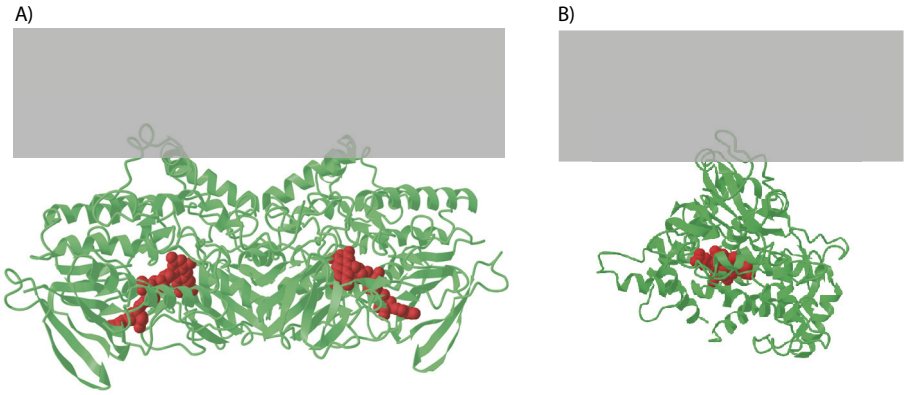


Figure 12.6 Molecular models of (A) a flavin monooxygenase and (B) a cytochrome P-450. The FAD and haem coenzymes are shown in red in each model. The grey blocks indicate the approximate positions of the smooth endoplasmic reticulum membrane. PDB files 6SEK and 3E6I, respectively.

One of the atoms of the dioxygen is incorporated into the xenobiotic to form a hydroxyl group, while the other oxygen atom is reduced to water. The $FADH_2$/NADPH supplies electrons to complete the reaction:

$$\text{R-H} + \text{O}_2 + \text{FADH}_2/\text{NADPH} + \text{H}^+ \xrightarrow{\text{MFO activity}} \text{R-OH} + \text{FAD/NADP}^+ + \text{H}_2\text{O}$$

where R-H represents the xenobiotic.

Over 50 different members of the cytochrome P-450 enzyme family are found in the SER of hepatocytes. This enzyme system can also convert unsaturated compounds to epoxides, which are substrates for epoxide hydrolase activity that catalyses their conversion to a glycol (*Figure 12.7*).

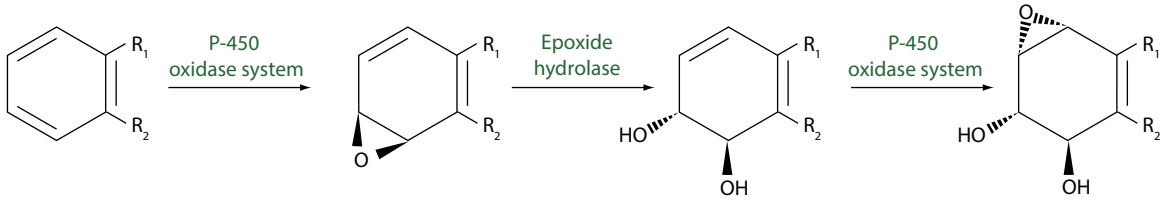


Figure 12.7 The combined actions of a P-450 oxidase and an epoxide hydroxylase in hydroxylating a hydrophobic xenobiotic (here shown as a saturated aromatic ring).

The net result of these enzyme activities is that the water solubility of the products is higher than that of the original poison or drug. The newly formed hydroxyl groups also constitute substrate attachment points for the actions of phase II enzymes, such as glucuronyltransferases of the SER membrane and sulphotransferase in the cytosol. These enzymes attach polar, that is, the hydrophilic, glucuronate or sulphate groups to the products of phase I reactions using UDP-glucuronate and 3′-phosphoadenosyl 5′-phosphosulphate (PAPS) as their respective donor substrates:

$$\text{R-OH} + \text{UDP-glucuronate} \xrightarrow{\text{Glucuronyltransferase}} \text{R-O-glucuronate} + \text{UDP}$$

$$\text{R-OH} + \text{PAPS} \xrightarrow{\text{Sulphotransferase}} \text{R-O-SO}_3^- + \text{PAP}$$

BOX 12.2 COLLATERAL DAMAGE BY THE P 450 SYSTEM

It is, unfortunately, the case that several poisons are rendered more toxic by their oxidation by the P-450 system. Indeed, a number of fat-soluble compounds that are relatively harmless, because they would normally accumulate in adipose tissue, are converted into potentially lethal materials because of the increase in water solubility. Thus, a number of indirectly acting carcinogens, for example benzo[a]pyrene found in cigarette smoke and the toxin aflatoxin B1 from *Aspergillus flavus*, are converted to products that are highly carcinogenic because they can react with guanine bases of DNA (*Figure 12.8* and *Chapter 17*).

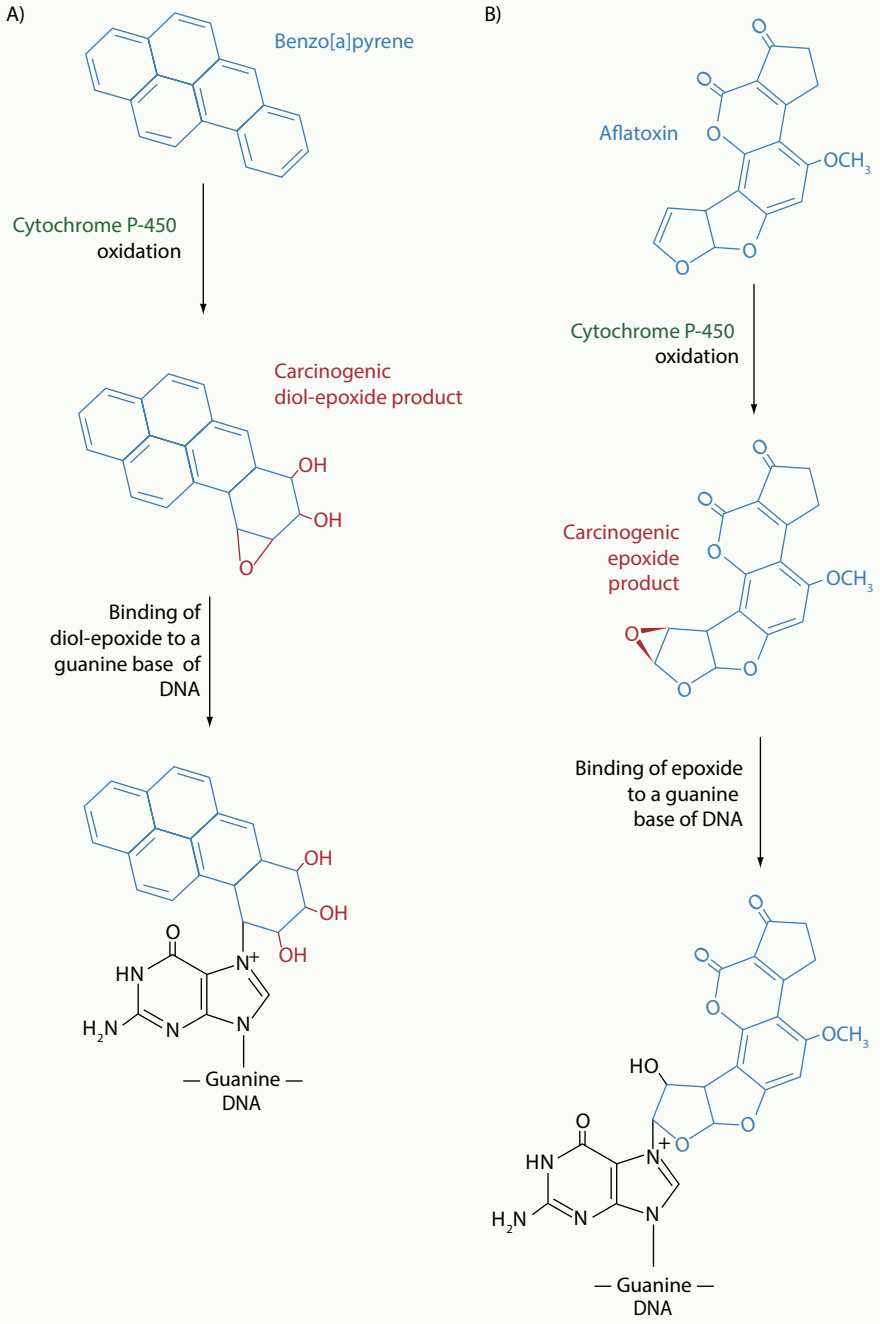


Figure 12.8 The oxidation of (A) benzo[a]pyrene and (B) aflatoxin by the cytochrome P-450 system to reactive chemicals that can bind with guanine bases of DNA and so induce carcinogenesis.

The increased water solubility of the final products facilitates their excretion by the kidneys, although small amounts are also lost in bile to the faeces (*Chapter 11*).

12.4 SYMPTOMATIC POISONING

In symptomatic patients, a diagnosis of poisoning is usually made on the basis of clinical findings. However, in all cases of suspected poisoning, several biochemical investigations should be made (*Table 12.1*). The roles of the clinical laboratory in chemical toxicology in identifying a suspected poison in blood or urine samples and in monitoring vital functions cannot be overestimated. The clinical management of patients is directed towards supporting their vital functions and attempting to remove the poison from the body as rapidly as possible. Unfortunately, few of the signs or symptoms that patients present are specific for any one type of drug or poison. Also, patients may be comatose and unable to give relevant information. A drug screen on a urine specimen may be carried out, but this will only indicate that a drug has been ingested and may give no indication of the severity of the overdose. Drug abuse is an ever-growing problem, particularly in Western countries, and is a significant cause of mortalities. For example, in England and Wales, there were over 5400 deaths related to drug poisoning in 2023. In 2022, drug overdoses were responsible for 105 452 deaths in the USA.

Investigation	Purpose
Monitor plasma urea, electrolytes and creatinine concentrations	Assess renal function (*Chapter 8*)
Monitor plasma osmolality	An osmolality gap can indicate the presence of foreign substances in the serum (*Chapter 8*)
Monitor blood gases	Assess acid–base status (*Chapter 9*)
Monitor liver functions	Assess liver viability (*Chapter 11*)
Monitor blood glucose concentration	To detect hypo- and hyperglycaemia (*Chapter 7*)

Table 12.1 Biochemical investigations of suspected poisoning. ⏎

12.5 COMMON POISONS

Numerous xenobiotics (drugs, pesticides, environmental pollutants, industrial chemicals and food additives) are potentially capable of perturbing the biochemistry and physiology of the body, either directly or after being metabolically transformed. However, the most frequent poisons encountered in emergency toxicology include:

- paracetamol (acetaminophen)
- aspirin
- alcohols (ethanol, methanol, ethylene glycol)
- barbiturates (although this is now mainly historical)
- carbon monoxide
- paraquat
- several metals

This chapter will concentrate only on those most likely to be encountered in clinical practice.

Paracetamol

Paracetamol (*Figure 12.9*) is used as an analgesic, that is, to relieve pain and fever. It is safe when taken at recommended doses but toxic if overdosed. It is the most common cause of poisoning admissions to hospital due to its wide availability. In the UK, paracetamol overdoses caused 227 deaths in 2021.

Paracetamol is rapidly absorbed from the stomach and upper intestine. The major part is metabolised by conjugation with glucuronide (approximately 60%) or sulphate (about 30%) in the liver as described in *Section 12.3* to form nontoxic metabolites. However, approximately 10% is metabolised by cytochrome P-450 to produce the highly reactive intermediate *N*-acetyl-*p*-benzoquinoneimine (NAPQI). Hepatocytes can metabolise NAPQI by forming conjugates with the tripeptide glutathione (see *Chapters 13* and *18*)

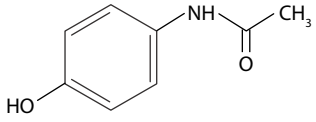


Figure 12.9 Paracetamol (called acetaminophen in the USA). ⏎

to produce a nontoxic mercapturic acid. However, when an overdose of paracetamol is ingested, liver detoxification systems may become saturated and the large amount of NAPQI produced exhausts the limited stores of glutathione. As a consequence, NAPQI binds to sulphhydryl groups of hepatocyte proteins, forming irreversible complexes that result in acute hepatic necrosis, that is, cell death (*Figure 12.10*). Paracetamol is also metabolised in renal tubule cells and, in an overdose, tubular necrosis may occur. In the presence of hepatic damage, there is usually only a small amount of renal damage, but occasionally this may be the major presenting feature of paracetamol poisoning.

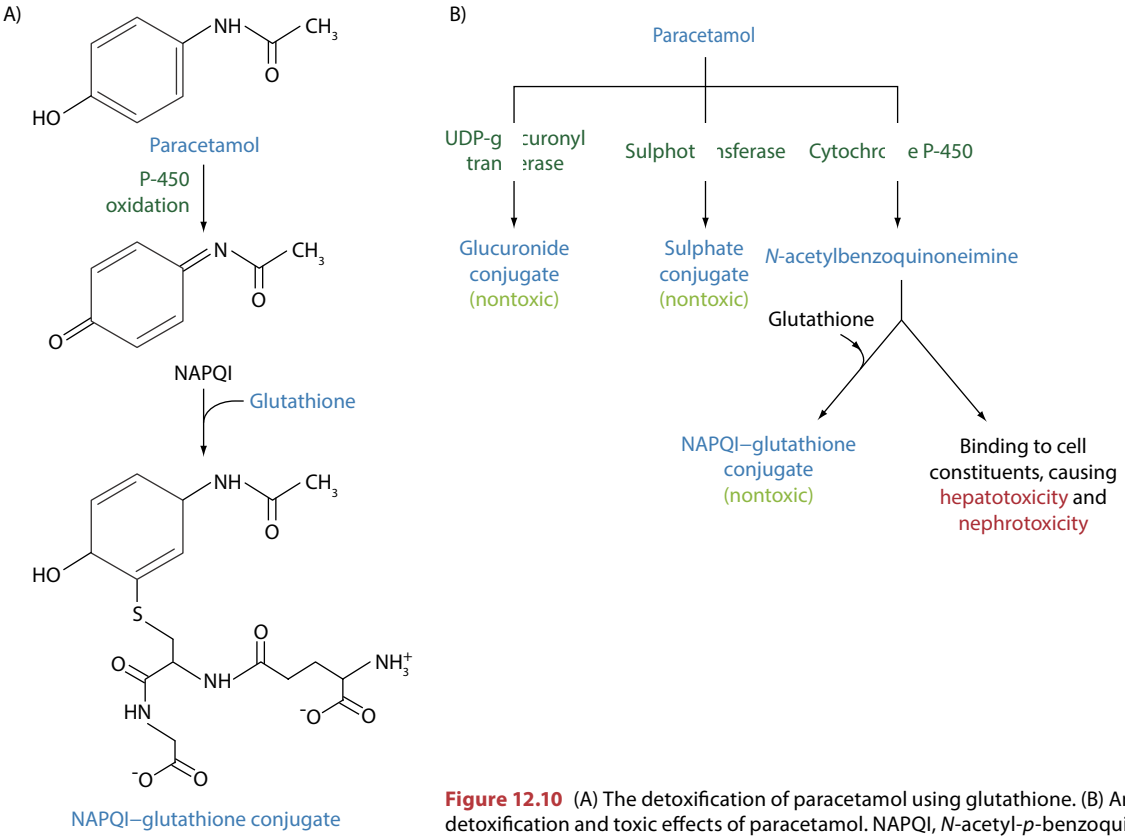


Figure 12.10 (A) The detoxification of paracetamol using glutathione. (B) An overview of the detoxification and toxic effects of paracetamol. NAPQI, *N*-acetyl-*p*-benzoquinoneimine.

Alcohol in chronic alcoholics and the phenobarbitone or phenytoin used to treat epilepsy may induce the biosynthesis of cytochrome P-450 enzymes and cause increased production of NAPQI. As a consequence, hepatotoxicity may occur following a relatively small overdose of paracetamol in such patients. Severe toxicity is also more likely in people whose intracellular stores of glutathione are depleted as a result of starvation or protein malnutrition (*Chapter 10*).

Toxic doses of paracetamol

A dose of 15 g for adults and 4 g for children is normally sufficient to cause hepatotoxicity. It has been suggested that this amount of paracetamol depletes liver glutathione concentrations by 70% in a 70 kg man. However, there is wide variation in the metabolic handling of paracetamol and large overdoses of over 50 g have been known to have little effect in some patients. The incidence of hepatotoxicity in children is significantly lower at concentrations of paracetamol in blood that would be potentially toxic in adults. Animal studies have suggested that turnover of glutathione is age dependent: younger animals can tolerate larger doses of paracetamol.

Often, estimates of the amount of paracetamol ingested, for example by a potential suicide, are unreliable. Thus, predictions of hepatotoxicity should be made only on the basis of serum concentrations. In general, concentrations of paracetamol greater than 300 mg/L cause serious liver damage four hours after ingestion, whereas values below 120 mg/L show no toxicity.

Clinical features of paracetamol poisoning

The signs and symptoms of paracetamol overdose are insidious, especially in the earlier stages. The clinical features can be divided into three phases, with a fourth occurring if the person survives, as described in *Table 12.2*.

Phase	Duration	Clinical features
I	0.5–24 hours	Loss of appetite and anorexia (*Chapter 10*) Nausea and vomiting General malaise Patient appears normal
II	24–48 hours	Less severe symptoms Abnormal blood chemistry with increases in liver enzyme activities and bilirubin Deterioration in renal functions but blood urea concentration remains low given the decrease in liver function (*Chapters 8*, *10* and *11*)
III	72–96 hours	Coagulation defects Jaundice and renal failure Reappearance of nausea and vomiting Signs of hepatic necrosis Death due to hepatic failure
IV	4–14 days	Hepatic and renal functions return to normal if patient survives phase III

Table 12.2 Phases of paracetamol poisoning. ⏎

Laboratory investigations of paracetamol poisoning

Estimating the concentration of paracetamol in plasma is useful for assessing the probability of patients developing hepatotoxicity. A nomogram (*Figure 12.11*) is available for paracetamol poisoning but should only be used when the size of the overdose and the approximate time of ingestion are known. Blood samples for paracetamol determination should be drawn at least four hours postdose to allow for its complete absorption and the plasma concentration to peak. The concentration of paracetamol in plasma can be used as a guide to patient management. Other tests that may be useful are determining the activities of liver transaminases and measuring serum bilirubin to monitor liver functions (*Chapter 11*), and determining serum creatinine concentrations to assess renal

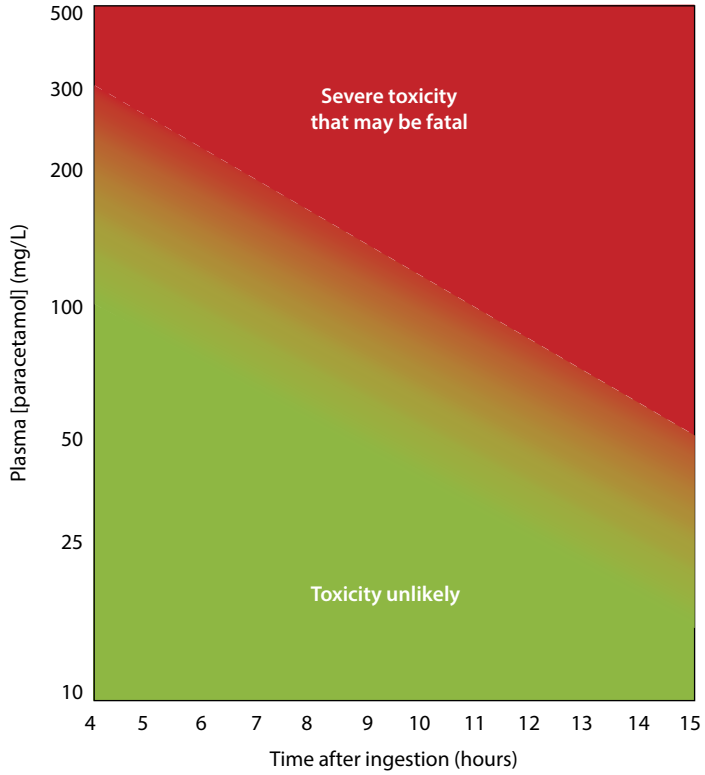


Figure 12.11 A nomogram to assess paracetamol toxicity. ⏎

function (*Chapter 8*). Paracetamol poisoning causes an increased prothrombin time (*Chapter 13*), the time taken for blood clotting to occur, in a sample of blood to which Ca^{2+} and thromboplastin have been added.

Management of paracetamol poisoning

General procedures include the administration of activated charcoal. This binds paracetamol and can reduce its absorption by the GIT if given in the first hour of an overdose. Gastric lavage may also be used in patients who have ingested large amounts of paracetamol and present within an hour of ingestion. Antidotes to paracetamol poisoning include methionine and *N*-acetylcysteine (*Figure 12.12*). Both promote the biosynthesis of glutathione in the liver, increasing its capacity to detoxify the active metabolite. Methionine may be given orally at 2.5 g every 4 hours for 12 hours in early or uncomplicated cases. *N*-acetylcysteine is administered parenterally and is more appropriate for patients who present late or are comatose or vomiting. *N*-acetylcysteine can act as a glutathione substitute and enhances conjugation with sulphate (see *Section 12.3*). It also limits liver damage by reducing inflammation and improving the microcirculation in the liver. Treatment gives maximal benefit if started within 10 hours of ingestion, although it may still be beneficial for up to 24–30 hours.

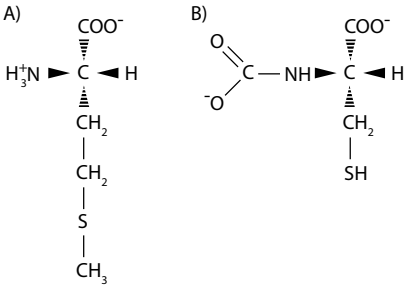


Figure 12.12 (A) Methionine and (B) *N*-acetylcysteine.

Aspirin

Aspirin (acetylsalicylate) is hydrolysed in the body to salicylate, the active form of the drug (*Figure 12.13*), which has analgesic, antipyretic and anti-inflammatory properties. Salicylate is eliminated from the body following conjugation reactions with glycine to produce salicyluric acid and, to a lesser extent, with glucuronide to form phenol and acylglucuronides. A small amount is hydroxylated to gentisic acid. In an overdose, these pathways may become saturated, and a large proportion of salicylate may be excreted unchanged in urine.

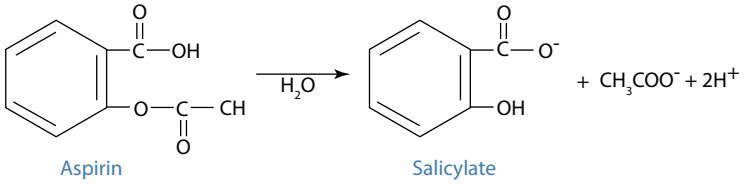


Figure 12.13 The conversion of aspirin to salicylate.

Like paracetamol, aspirin is widely and easily available and poisoning by aspirin overdose is therefore relatively common. Salicylate toxicity is due to several effects (*Figure 12.14*). An increase in the concentration of salicylate in the brain stimulates the respiratory centre, leading to hyperventilation, which can result in dehydration. In addition, hyperventilation causes a decline in the PCO_2 and a rise in pH, that is, a respiratory alkalosis. As PCO_2 declines, the pH increases and less H_2CO_3 is formed. The kidneys compensate by excreting more hydrogen carbonate (HCO_3^-) and K^+ whilst retaining H^+. These activities, rather than helping correct the respiratory alkalosis, contribute towards a latent metabolic acidosis (*Chapter 9*). Salicylate also uncouples oxidative phosphorylation, decreasing ATP formation and increasing heat production, with sweating and hyperpyrexia (compare with DNP; see *Box 12.1*), which further contributes to fluid loss and dehydration. The decline in the amount of ATP stimulates glycolysis, and therefore pyruvate and lactate accumulate. An increase in glycogenolysis (*Margin Note 12.2*) provides the required glucose for the increased glycolysis. Eventually, there may be depletion of glycogen reserves.

Toxic doses of aspirin

Common therapeutic concentrations of salicylate are 50 mg/L, although they can be as high as 250 mg/L in the serum of patients with rheumatoid arthritis (*Chapter 5*). The potentially lethal dose of aspirin in adults is 24–30 g, but death can occur in children under 18 months from as little as 300 mg. Signs of salicylate toxicity occur when concentrations are greater than 300 mg/L. In severe cases of aspirin poisoning, the concentrations in serum can be as high as 1000 mg/L.

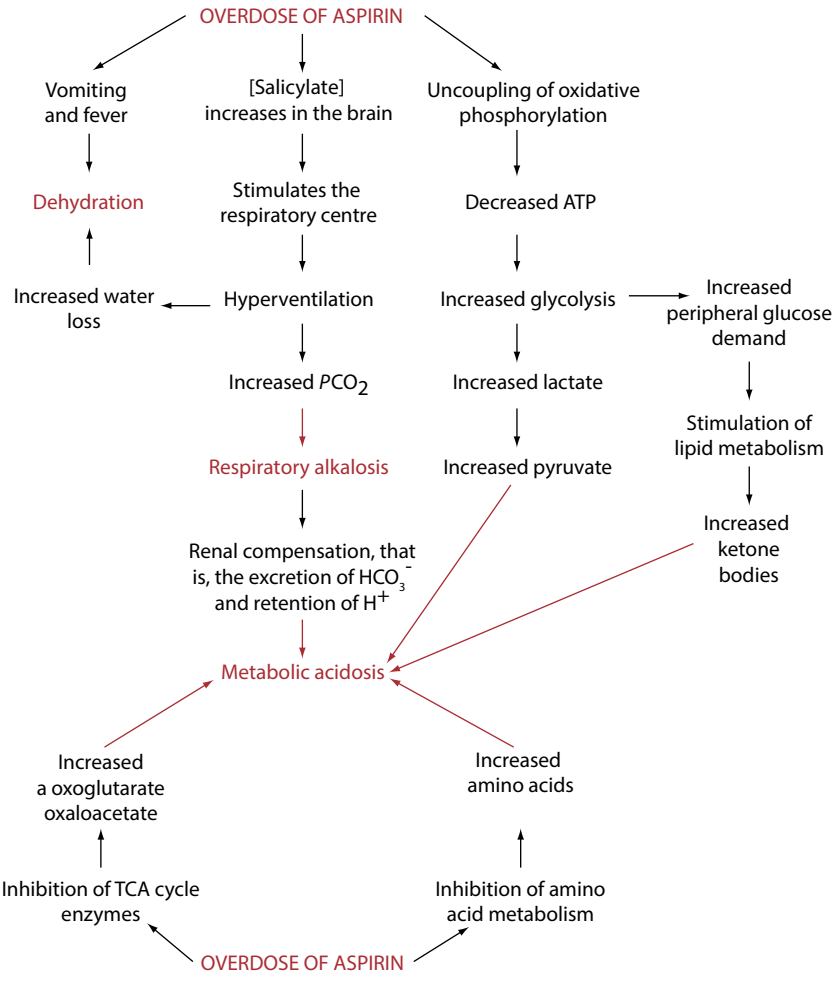


Figure 12.14 An overview of the toxic effects of aspirin.

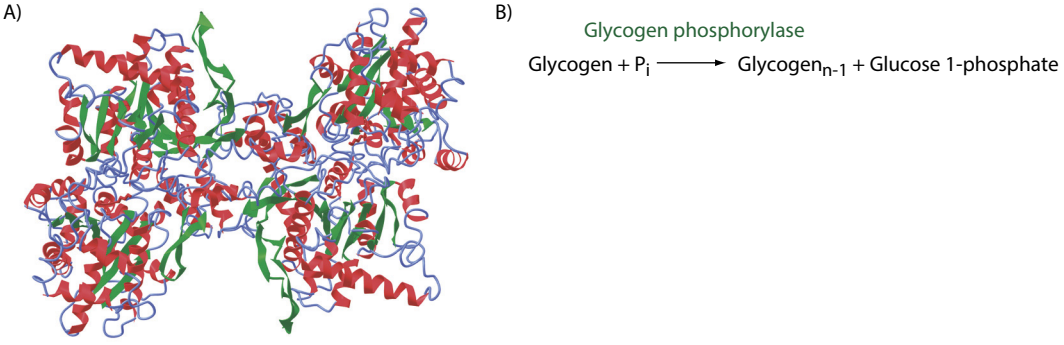


Figure 12.15 (A) Molecular model of glycogen phosphorylase. PDB file 3CEH. (B) The action of glycogen phosphorylase in releasing glucose from stores of glycogen.

Clinical features of aspirin poisoning

Numerous symptoms are associated with aspirin poisoning, including confusion and/or unconsciousness, hyperventilation, nausea, sweating, tinnitus (buzzing/high-pitched noise in ear), vomiting and a severe loss of fluid.

Laboratory investigations of aspirin poisoning

The concentration of salicylate in plasma should be measured on presentation and every four to six hours until it has fallen below the toxic range. This is necessary, as salicylates precipitate in acid conditions and may therefore deposit in large amounts in the

Margin Note 12.3 Glycogenolysis

Glycogenolysis is the breakdown of glycogen reserves to yield glucose 1-phosphate in a reaction catalysed by glycogen phosphorylase (*Figure 12.15*). Glucose 1-phosphate can be oxidised to give ATP for use by the cell or, in liver cells particularly, hydrolysed to form glucose that can be released into the bloodstream to maintain blood sugar concentrations.

stomach. The consequence is delayed absorption, and this means that the concentration in plasma may continue to rise for many hours following a severe overdose.

Management of aspirin poisoning

A variety of measures can alleviate salicylate poisoning. These are aimed at decreasing its absorption, increasing its rate of elimination and correcting the associated acid–base and electrolyte disturbances. These measures include gastric lavage for up to 24 hours after ingestion, as salicylate may remain unabsorbed in the GIT for long periods. Patients may also be given 50 g of activated charcoal followed by 25 g every four hours. The ionisation state of salicylate affects its reabsorption by the kidneys. If the provisional urine is acidic, salicylic acid is unionised and is filtered at the glomerulus but reabsorbed from the tubules (*Chapter 8*); if the urine is alkaline, then salicylate predominates and its tubular reabsorption is reduced and more is lost from body. Hence, patients are infused with sodium hydrogen carbonate ($NaHCO_3$) to increase the pH of provisional urine and promote the loss of salicylate. Fluid replacement, corrections of the acid–base and electrolyte imbalances (especially the hypokalaemia), and improving the hypoglycaemia are all required. Haemodialysis may be necessary in severe cases of poisoning, such as when plasma concentrations of salicylate exceed 800 mg/L. As well as removing the salicylate, haemodialysis also corrects the acid–base and electrolyte imbalances. The technique requires a dialyser (*Figure 12.16*). Blood from a patient's artery is circulated through the dialyser on one side of a semipermeable membrane while a solution of normal electrolytic composition circulates on the other side. Waste products, poisons (including salicylate) and small molecules cross the membrane and are therefore removed, and the dialysed blood is returned to the body via a catheter in a vein.

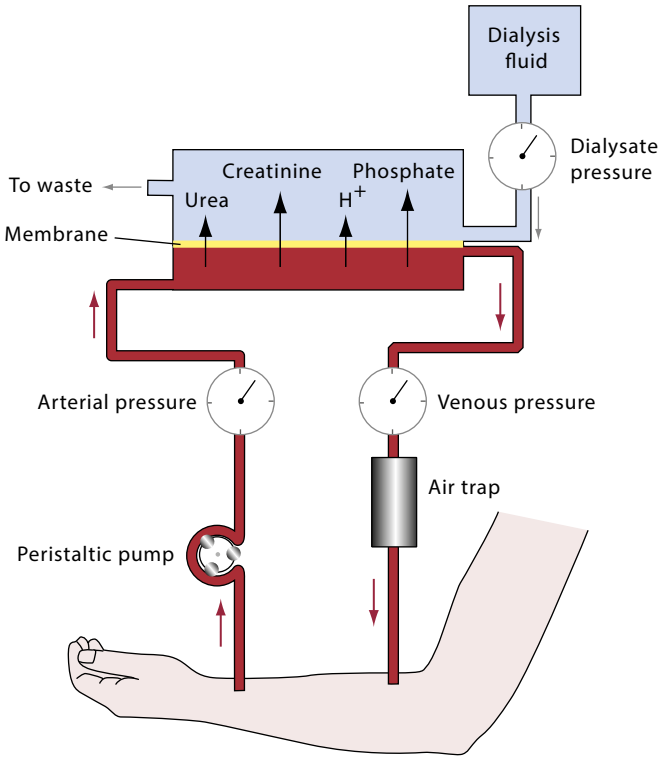


Figure 12.16 A schematic illustrating the procedure of haemodialysis.

Ethanol

Ethanol is an addictive drug and its abuse can lead to dependency and alcoholism. Its abuse is increasingly common in the developed world. The toxic effects of chronic alcohol abuse on the liver, brain and GIT are widely known. Ethanol can also modify the effects of other drugs; for example, it inhibits the hydroxylation of barbiturates (see later) by the P-450 system, inhibiting their excretion by the kidneys.

The metabolism of ethanol occurs mainly in the liver by one of two mechanisms. Normally, only small amounts are degraded by the P-450 system (*Figure 12.17*), although this oxidation can become of greater importance because the P-450 system is induced

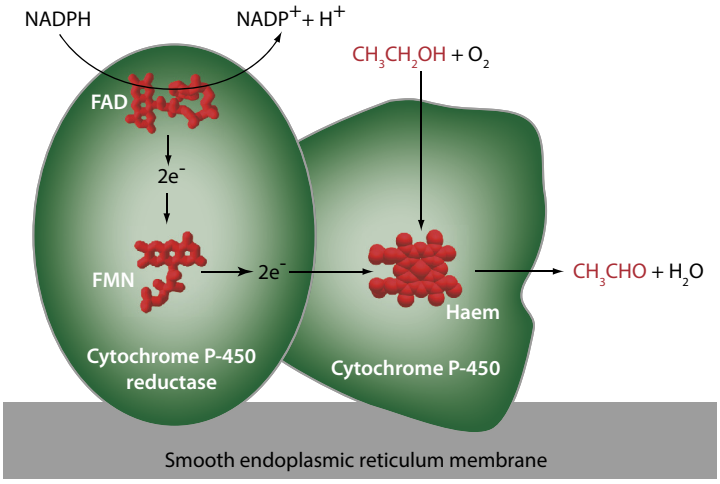


Figure 12.17 The role of P-450 in the detoxification of ethanol.

by chronic alcohol consumption. The major catabolic pathway is to oxidise ethanol to the corresponding aldehyde, ethanal, in a reaction catalysed by alcohol dehydrogenase (ADH). The ethanal may subsequently be oxidised to ethanoate (acetate) by aldehyde dehydrogenase (ALDH). These reactions are summarised in *Figure 12.18*.

Margin Note 12.4 Ethanol dehydrogenase and ethnicity

The effectiveness (specific activity) of alcohol dehydrogenase varies between different ethnic groups. Thus, the (often undesirable) effects of alcohol drinking can appear after widely varied times, irrespective of the quantity consumed.

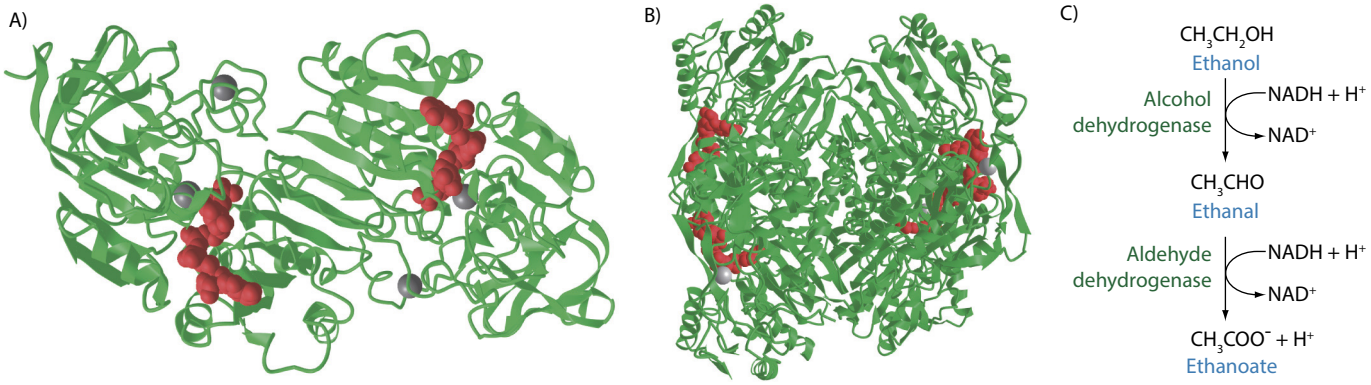


Figure 12.18 Molecular models of (A) an alcohol dehydrogenase (the bound NAD^+ is shown in red and the grey spheres represent Zn^{2+}) and (B) an aldehyde dehydrogenase (the bound NAD^+ is shown in red and the grey spheres represent Mg^{2+}); PDB files 1HSO and 1CW3 respectively. (C) The reactions catalysed by alcohol and aldehyde dehydrogenases.

Much of the ethanoate made from ethanol escapes into the blood and can result in an acidosis (*Chapter 9*). The activities of both enzymes increases the NADH/NAD^+ ratio, that is, alcohol consumption leads to the accumulation of NADH with consequent severe effects. The increased NADH inhibits fatty acid oxidation and stimulates the biosynthesis of triacylglycerols in the liver, producing a fatty liver. The oxidation of lactate to pyruvate is also inhibited, slowing gluconeogenesis (*Margin Note 12.5*). The increased lactate exacerbates the acidosis, and the decreased gluconeogenesis may cause hypoglycaemia.

The capacity of liver mitochondria to oxidise acetate to CO_2 is limited because the activation of ethanoate/acetate to acetyl CoA requires ATP:

$$\text{Acetate} + \text{coenzyme A} + \text{ATP} \rightarrow \text{acetyl CoA} + \text{AMP} + \text{PP}_i$$

However, the supply of ATP is decreased because glycolysis, which requires free NAD^+, is reduced and because the use of acetyl CoA by the TCA cycle is blocked, since NADH inhibits the regulatory enzymes isocitrate dehydrogenase and 2-oxoglutarate dehydrogenase. Acetyl CoA is converted to ketone bodies (*Chapter 7*), which are released into the blood, intensifying the acidosis caused by acetate and lactate. Ethanal also accumulates, and this extremely reactive compound can bind to liver proteins, impairing their functions and causing severe damage and the death of hepatocytes.

Liver damage from excessive alcohol consumption occurs in three stages. The first stage is the formation of the fatty liver described above. This condition, in the absence

Margin Note 12.5 Gluconeogenesis

Gluconeogenesis is the biosynthesis of glucose using materials that are not carbohydrates as precursors. These include pyruvate, the end product of glycolysis; oxaloacetate, an intermediate of the TCA cycle; and dihydroxyacetone phosphate, which can be made from glycerol obtained from the hydrolysis of triacylglycerols.

of other complications, is readily reversible within four to six weeks if alcohol is avoided. The second stage is the occurrence of alcoholic hepatitis when groups of liver cells die. This leads to inflammation and can be fatal. In the third stage, the patient may develop cirrhosis, a condition seen in 10–15% of alcoholics (*Figure 12.19*). Approximately 50% of all cases of cirrhosis are due to alcoholic liver disease. Cirrhosis occurs when fibrous structures and scar tissue are produced around the dead cells. This impairs many of the biochemical functions of the liver; for example, a cirrhotic liver cannot convert ammonia to urea and the concentration of ammonia in the blood increases. Ammonia is toxic to the nervous system and can cause coma and death. Ethanol also directly affects the central nervous system (CNS). For example, it enhances the inhibitory effects of γ aminobutyric acid (GABA) at $GABA_A$ receptors and the functions of some 5-hydroxytryptamine receptors, in addition to numerous other membrane proteins. The resulting depressive effects of these activities are well known.

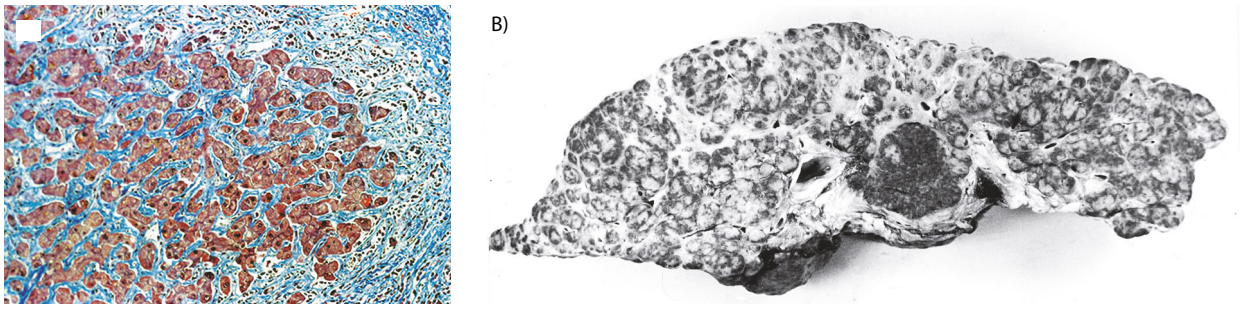


Figure 12.19 Cirrhosis in alcohol-related liver disease. (A) Note how alcohol abuse has led to diffuse scarring of the liver, which has regenerated at numerous points to form nodules. The sinusoidal fibrosis is well established in this small regenerative nodule where the regenerating cells are darkly stained compared with the lighter fibrous tissue. As the nodules become larger, the fibrous tissue around the outside becomes compressed, giving the characteristic hobnail appearance to the liver as shown in (B), which is a transected cirrhotic liver from an alcoholic. The extensive damage in the form of nodules of varying sizes separated by bands of fibres is apparent. (A), courtesy of Professor F.A. Mitros, University of Iowa, USA; (B), courtesy of Dr N.P. Kennedy, Trinity College Dublin, Republic of Ireland. ⏎

Chronic liver damage invariably results in malnutrition, partly due to malabsorption and partly because the metabolism of nutrients by the liver is defective. However, many alcoholics are malnourished because of dietary inadequacies. While ethanol supplies most of their required energy, they may be ingesting insufficient amounts of other nutrients, particularly proteins and vitamins (*Chapter 10*). In addition, chronic alcohol ingestion can damage the mucosal lining of the GIT and pancreas, as well as the liver (*Chapter 11*). Alcoholics require increased amounts of vitamins and some trace elements because of the metabolic load experienced and their increased excretion. For example, niacin, although not strictly a vitamin since it can be formed, albeit very inefficiently, from tryptophan (*Chapter 10*), is necessary to form the coenzymes NAD^+ and $NADP^+$. A borderline deficiency of niacin leads to **glossitis** (redness) of the tongue, while a pronounced deficiency results in pellagra, with dermatitis, diarrhoea and dementia. In developed countries, pellagra is rarely encountered other than in the severe malabsorption problems of alcoholics; hence, dietary supplies of niacin and tryptophan are required. A severe zinc deficiency (*Chapter 10*) also occurs, primarily in alcoholics, especially those suffering from cirrhosis. Heavy drinkers frequently suffer from GIT varicose veins (*Chapter 14*) and diarrhoea caused by a variety of factors, including ethanol-exacerbated lactase deficiency and interference with normal peristalsis. Steatorrhoea is also common, due to deficiencies in folate and bile salts in the GIT (*Chapter 11*).

The obvious treatment for alcoholic liver disease and cirrhosis is abstinence. Given the addictive nature of ethanol, this may be difficult to maintain. Ethanal can also escape from the liver and react with blood proteins, forming stable adducts. These can provide useful clinical markers of the past drinking activity of a patient. Unfortunately, patients presenting with severe alcoholic hepatitis have a high mortality, even when abstinence is successful. Although a number of pharmacological treatments have been attempted as therapies, none has been successful. However, supportive nutritional management in cases of acute and chronic liver disease is essential to achieve electrolyte, vitamin and protein replenishments. The only cure for advanced cirrhosis is a liver transplant (*Chapter 6*).

The major enzyme involved in ethanol metabolism, ADH, has a broad specificity and can catalyse the oxidation of a range of alcohols. This allows ethanol to be used as a competitive inhibitor of the enzyme in the treatment of poisoning by other alcohols, such as methanol and ethylene glycol.

Methanol

Methanol is widely used as a solvent and as antifreeze. It is a potent poison: as little as 10 mL causes blindness, while 30 mL can be fatal. Toxicity can come from ingestion, inhalation and skin exposure. Mass poisonings have occurred from drinking alcoholic beverages made with contaminated ethanol and from accidental exposure. The major cause of methanol toxicity is its initial oxidation to methanal (formaldehyde), which is then converted to methanoate (formate). The first step in the pathway is catalysed by ADH and the production of formate is the result of several enzyme activities.

While both methanal/formaldehyde and methanoate/formate are toxic, the former has a short metabolic half-life, whereas formate accumulates, since its metabolism to CO_2 is slow in humans. This leads to a metabolic acidosis. The symptoms of methanol poisoning are an initial mild inebriation and drowsiness. Visual disturbances such as blurred vision, diminished visual acuity and dilated pupils occur after about six hours. Within 8–36 hours, nausea, vomiting, abdominal pain, headaches and possibly coma occur. Treatment for methanol poisoning is, firstly, to administer ethanol. This blocks methanol metabolism, since ADH has a greater affinity for ethanol than methanol. Secondly, $NaHCO_3$ is given intravenously to correct the metabolic acidosis.

Ethylene glycol

The alcohol ethylene glycol is a common component of antifreezes, paints, polishes and cosmetics. It is a poison with a minimum lethal dose of approximately 100 mL. Its toxicity is due to its oxidation to oxalate, which humans cannot rapidly excrete, and also results in an excess of NADH resulting in lactate production (*Figure 12.20*). The oxalate and lactate can result in a metabolic acidosis, while the organic acids produced by the breakdown of ethylene glycol inhibit several metabolic processes, including oxidative phosphorylation. The production of oxalate results in its deposition as insoluble calcium oxalate in renal tubules and the brain. These biochemical events reduce the concentrations of HCO_3^- (hydrogen carbonate, 'bicarbonate') and Ca^{2+}, but increase that of K^+ in plasma. Crystals of calcium oxalate, blood and protein may all leak into the urine, giving rise to three distinct clinical phases. Within 30 minutes to 12 hours, convulsions, coma, intoxication, nausea, nystagmus, papilloedema, depressed reflexes, myoclonic jerks, tetanic contractions, permanent optic atrophy and vomiting may occur. In 12–23 hours, congestive cardiac failure, hypertension, rapid breathing (tachypnoea) and heart rate (tachycardia), and pulmonary oedema may all present. In the following 24–72 hours, kidney damage with pain and acute renal tubular necrosis may feature. Death may occur within 24 hours due to damage to the CNS, or between 8 and12 days due to renal failure.

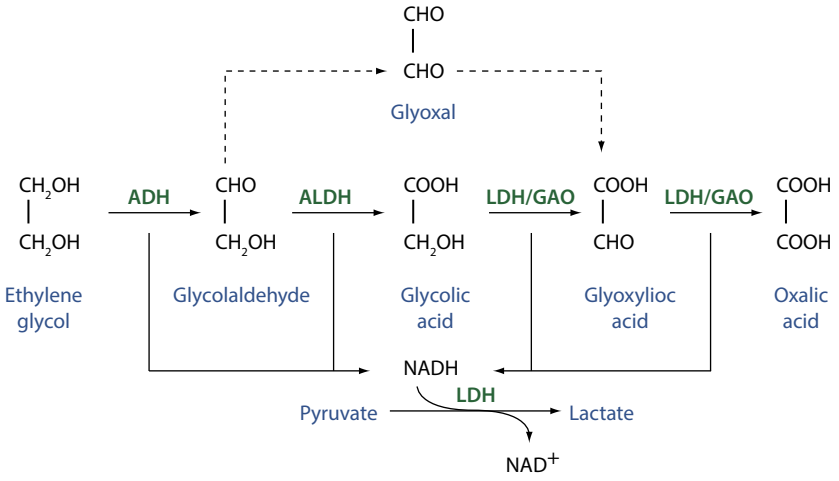


Figure 12.20 The formation of lactate during the catabolism of ethylene glycol. ADH, alcohol dehydrogenase; ALDH, aldehyde dehydrogenase; LDH, lactate dehydrogenase; GAO, glycolic acid oxidase.

Treatment of ethylene glycol poisoning is the application of gastric lavage to reduce its absorption, combined with supportive therapies for the shock and respiratory distress. Administration of ethanol is standard, since it competes effectively for the active site of ADH, inhibiting the catabolism of the absorbed ethylene glycol. Sodium hydrogen carbonate, administered intravenously, and calcium gluconate are used to correct the acidosis and hypocalcaemia, respectively. Dialysis is also used to remove ethylene glycol.

Barbiturates

Barbiturates are a group of drugs based on the parent compound barbituric acid (*Figure 12.21*). All are sedatives, that is, they depress certain activities of the CNS. Although they have well-established therapeutic uses, barbiturates can be toxic if taken as an overdose. They were once used as recreational drugs, for example in purple hearts, and this too led to accidental overdoses. Therapeutically, barbiturates have now largely been replaced by benzodiazepines, and barbiturate overdose is less likely to be encountered today, although 83 deaths related to barbiturate use occurred in England and Wales in 2019–2021.

Barbiturates induce drug dependence and this involves three distinct and independent components:

1. tolerance
2. physical dependence
3. compulsive abuse or psychic craving

Barbiturate dependence stimulates all three components to such an extent that it produces major problems for the individual user in addition to society at large. Once barbiturate dependence has developed, abrupt withdrawal from the drug induces a particularly unpleasant withdrawal syndrome. This is characterised by anxiety, convulsions of the grand mal type, increased respiration and pulse with a corresponding increased blood pressure, insomnia, loss of weight, tremors, vomiting, weakness and a psychosis resembling alcoholic delirium tremens. Thus, during withdrawal, it is appropriate to reduce the dosage gradually over an extended period.

The length of time that barbiturates act *in vivo* varies. Some, for example thiobarbitone, are short acting; others, such as pentobarbitone, act in the medium term; while the group including the best-known barbiturate, phenobarbitone (*Figure 12.21*), induce long-lasting effects. Barbiturates act by enhancing the effects of GABA by binding to different sites on the receptors of target neurons in the CNS from those for GABA itself. The blood–brain barrier (see *Figure 3.8*) prevents the easy entry of many substances into the brain compared with their uptake into other tissues. However, lipid-soluble substances, such as the barbiturate thiopentone, enter the brain quickly by passive diffusion. This rapid uptake allows thiopentone to exert its anaesthetic effects extremely quickly. In contrast, other barbiturates, such as phenobarbitone, are weak acids and so ionised at physiological pH (see, for example, *Margin Note 8.3*). This slows their entry into the CNS but produces a longer lasting effect. The life-threatening effects of barbiturate poisoning include depression of the centres in the CNS that control respiration and blood circulation. Some barbiturates, for example phenobarbitone, can also induce the biosynthesis of monooxygenase enzymes (P-450; *Section 12.3*) and so, by altering the rate or route of metabolism of other drugs, can change their toxicities.

There is no antidote for barbiturate poisoning. Primary care is to maintain a free airway, use artificial respiration if required and administer a forced alkaline diuresis. The final point refers to the use of a isotonic saline containing diuretic agents combined with hydrogen carbonate to increase the glomerular filtration rate and the production of urine.

The pH dependence of ionisation of many common barbiturates is exploited by infusions of large volumes of $NaHCO_3$ containing the osmotic diuretics urea or mannitol. This increases the pH of plasma relative to the cytoplasm of cells and increases the proportion of ionised barbiturate in the plasma, causing more of the unionised drug to diffuse out of the tissues, including the brain, into the plasma. This can promote a diuresis as large as 12 L in 24 hours. The ionised form is also excreted more rapidly, since the alkalinity of the provisional urine ensures it remains ionised and cannot cross the tubule wall back into the plasma.

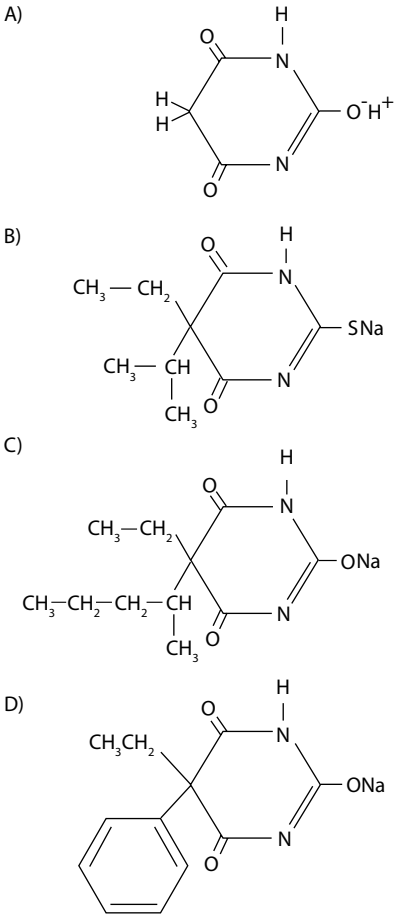


Figure 12.21 The barbiturates (A) barbituric acid, (B) thiobarbitone, (C) pentobarbitone and (D) phenobarbitone.

Carbon monoxide

Carbon monoxide (CO) is a poisonous gas. However, it is not an irritant and is odourless. Hence, it is insidious, and concentrations can build up, with the victim being unaware of any danger. Carbon monoxide is no longer present in domestic gas in the UK but is still a major cause of poisoning, both accidental and intentional, and results in the deaths of approximately 30 people in the UK annually. Sources of CO include domestic fires, ovens and boilers, coal gas, furnace gas, cigarette smoke, burning plastic and car exhausts, although catalytic converters have reduced the CO output from car petrol engines. Thus, traffic police officers, firefighters, those trapped in fires and some factory workers are all potentially at a greater risk than the general population. In the UK, the major cause of poisoning results from exposure to inefficient oxidation in engines and poorly maintained gas fires, ovens and boilers, especially where ventilation is inadequate.

The mechanism of CO poisoning is well understood at the biochemical level. The gas is absorbed rapidly through the lungs and binds to the Fe^{2+} of haemoglobin at the same site as dioxygen, but about 240 times more strongly. The product of CO binding is carboxyhaemoglobin (*Figure 12.22*), which prevents the efficient distribution of oxygen to the tissues. Carbon monoxide is potentially extremely poisonous at low concentrations. The main target organs of CO poisoning are the heart and brain, which extensively utilise aerobic metabolic pathways, and so their abilities to sustain an oxygen debt are relatively poor.

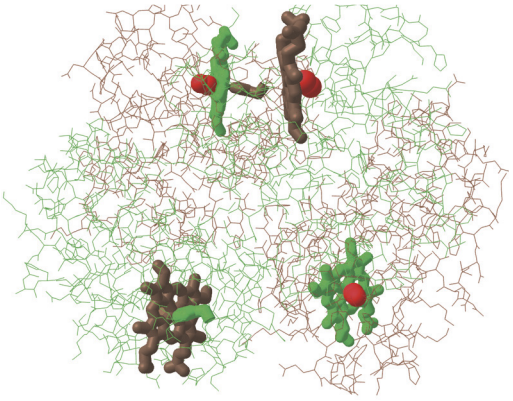


Figure 12.22 Molecular model of carboxyhaemoglobin. The bound carbon monoxide molecules are shown in red. PDB file 1AJ9. Compare to *Figure 13.9.*

Air contains 21% oxygen, but approximately only 0.1% CO will saturate 50% of the haemoglobin. Concentrations of 60% carboxyhaemoglobin in the blood are usually fatal, even if maintained for only a few minutes. A concentration of 20% carboxyhaemoglobin may not present obvious symptoms, but the ability to perform tasks can be impaired. At 20–30%, the victim may have a headache, with raised pulse, dulling of the senses and feelings of weariness. Concentrations of 30–40% accentuate these symptoms and decrease the blood pressure so that exertions may lead to faintness. At 40–60% and above, the victim becomes unconscious and will suffer convulsions. Other clinical features include nausea, pink skin, vomiting, loss of hearing, hyperpyrexia, hyperventilation, a decrease in light sensitivity, renal failure and acidosis. Death is due to brain tissue hypoxia, although cardiac arrythmias and heart and respiratory failures may also occur (*Chapter 14*). The duration of exposure is also a factor, since hypoxic cell death is not instantaneous. Also, some individuals, for example those with anaemia (*Chapter 13*), are more sensitive to CO poisoning than healthy people.

The treatment of CO poisoning involves removing its source and supplying the victim with fresh air or oxygen. The use of 100% oxygen at 2.5×10^5 Pa pressure, that is, hyperbaric oxygen, increases the rate of dissociation of carboxyhaemoglobin and reduces its plasma half-life from 250 minutes in a patient breathing air to 23 minutes. Adding carbon dioxide can be useful, as this reduces the half-life to 12 minutes at normal pressure.

Paraquat

Paraquat (*Figure 12.23*) is used widely as a weed killer, hence is often found in the home (or garden shed). Unfortunately, paraquat has caused many hundreds of deaths by both accidental and deliberate poisoning, the latter including both suicides and homicides.

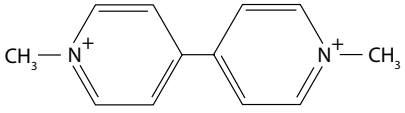


Figure 12.23 Paraquat.

Paraquat is a local skin irritant, causing inflammation, although poisoning usually follows oral ingestion. The toxic effects are dose related, and small amounts cause minimal damage, which is reversible. Fatal doses cause death within several days or weeks, with extreme abdominal pain, diarrhoea and vomiting. The major target organs are the lungs, with larger doses causing alveolar oedema, resulting in destruction and fibrosis of lung tissues if the patient survives longer than a few days. Pulmonary fibrosis and respiratory failure can, however, develop up to six weeks after ingestion. The kidneys, heart and liver may also be damaged, but the lungs are particularly susceptible to paraquat poisoning because alveolar epithelial cells actively accumulate paraquat to toxic concentrations. Furthermore, the presence of large concentrations of oxygen in these organs exacerbates the morbid effects. In alveolar epithelial cells, paraquat is reduced by electron donors, such as NADPH, to a stable reduced form (*Figure 12.24*). Given the prevailing aerobic conditions, the reduced paraquat transfers an electron to dioxygen to form a superoxide radical. Superoxide dismutase (SOD) activity removes superoxide radicals by catalysing their conversion to hydrogen peroxide. The hydrogen peroxide

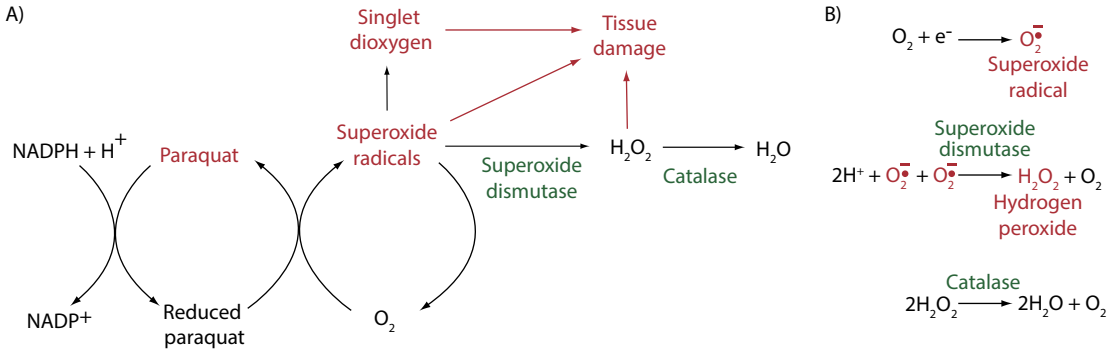


Figure 12.24 (A) An overview of the detoxification of paraquat and its toxic products. (B) The formation of superoxide radicals and the catalytic activities of superoxide dismutase (SOD) and catalase in degrading them to water and dioxygen.

may then be degraded to water and dioxygen by the action of catalase. *Figure 12.24* summarises these detoxification reactions. However, in aerobic conditions, the actions of SOD and catalase are likely to be overwhelmed, and superoxide radicals can accumulate and react with hydrogen peroxide to form highly toxic hydroxyl radicals, particularly in the presence of transition metal ions, such as those of iron and copper. This can lead to a variety of toxic effects, such as lipid peroxidation (*Figure 12.25*). Lipid peroxides can give rise to lipid radicals, which can oxidise membranes, leading to tissue damage and fibrosis. Lipid peroxides will also oxidise glutathione (*Figures 13.45* and *18.5*) and so further reduce the depleted amount of NADPH. This, in turn, reduces the ability of the alveolar cells to carry out essential biosynthetic repairs.

There is no antidote for paraquat poisoning, and once it has accumulated in the lungs little can be done to prevent its toxic effects. Treatment largely consists of trying to prevent absorption by the GIT by gastric lavage and by using fuller's earth as an adsorbent. Haemoperfusion may also be used to reduce the concentration of any paraquat already absorbed.

Metals

High concentrations of some metals are toxic. To diagnose metal toxicity, three features must be identified, namely:

1. a source of the toxic metal
2. signs and symptoms typical of toxicity by that metal
3. increased concentrations of the metal in body tissues

Metals that are commonly screened for toxicity include lithium, aluminium and the heavy metals lead, arsenic, cadmium and mercury.

Lithium

Lithium is widely used therapeutically to treat patients with certain psychiatric disorders. However, plasma concentrations of lithium greater than 1.5 mmol/L should be avoided and regular measurements of serum lithium concentrations are important in monitoring therapy. Lithium toxicity is associated with blurred vision, drowsiness, hypothyroidism, polyuria, tremors, tinnitus and, in severe cases, renal failure and coma.

Aluminium

Acute poisoning with aluminium is extremely rare. Indeed, aluminium compounds are clinically used for their antacid properties. The dietary intake of aluminium is 5–10 mg/day and this amount is removed completely by the kidneys. Unfortunately, patients with renal failure are susceptible to aluminium toxicity given their problems in its elimination. As the water used in dialysis may contain aluminium, which can enter the body through the dialysis membrane, the metal can build up to toxic concentrations and cause osteodystrophy and encephalopathy.

Aluminium toxicity is diagnosed by determining its concentration in plasma. Chronic toxicity occurs at concentrations above only 3 µmol/L, whereas 10 µmol/L can cause acute poisoning. Treatment is aimed mainly at prevention. When aluminium poisoning does occur, then its excretion may be enhanced using chelating agents, such as deferrioxamine mesylate (see *Figure 6.17*).

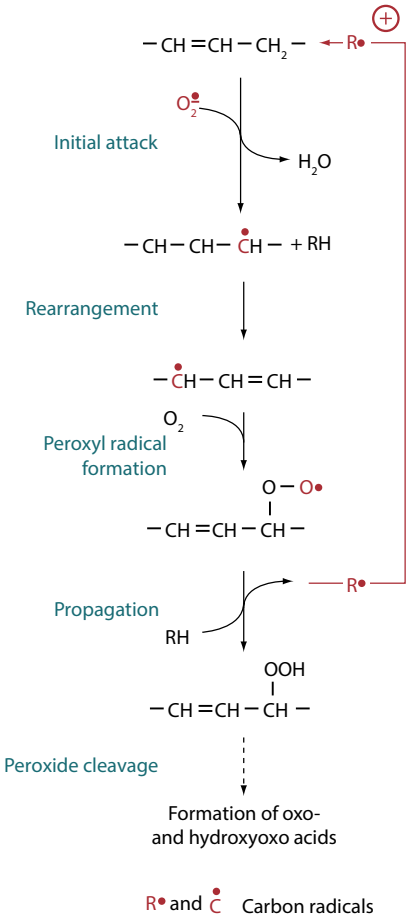


Figure 12.25 An overview of lipid peroxidation reactions. The peroxidation is initially started by a reactive radical. However, as carbon radicals are formed, the peroxidation becomes self-sustaining.

BOX 12.3 ALUMINIUM POISONING AND CAMELFORD, CORNWALL

If aluminium sulphate, $Al_2(SO_4)_3.18H_2O$ (*Figure 12.26*), is dissolved in water, it hydrolyses to form sulphuric acid and a gelatinous precipitate of aluminium hydroxide:

$$Al_2(SO_4)_3+6H_2O \rightarrow 3H_2SO_4+2Al(OH)_3\downarrow$$

An 8% solution of aluminium sulphate is commonly used in sewage treatment and water purification (see *Figure 2.2*) because it is a highly effective coagulant. It reacts with suspended organic matter to form aggregates of large flocs that rapidly settle, reducing the load of suspended solids and the turbidity of the water or sewage, thus improving their quality and lengthening the effective life of water and sewage filters.

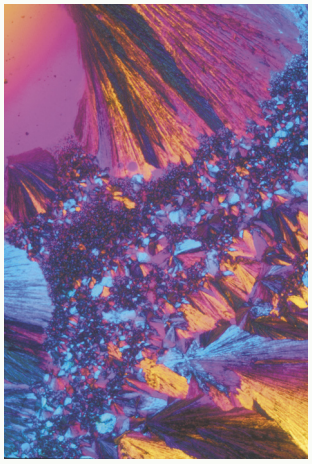


Figure 12.26 Hydrated aluminium sulphate as seen using polarised light.

The discharge of aluminium sulphate into aquatic environments should be minimised, as it can clog the gills of marine organisms. It is a weak acid and should be treated accordingly: contact with eyes, skin and clothing and repeated and prolonged exposure should be avoided. However, tests have shown it to be a minimal irritant and a nonirritant to rabbit eyes and skin, respectively. Its lethal dose 50 (LD_{50}) dose in rabbits, that is, the amount of aluminium required to kill 50% of a test population of the animals, is greater than 5 g/kg; hence, it is generally thought that its ingestion is relatively harmless.

Unfortunately, in 1988, a contractor accidentally dumped 20 tonnes of aluminium sulphate into the wrong tank at a water treatment plant that supplied domestic water to the 2500 residents of Camelford, a small town in Cornwall, UK. The mistake was later attributed to the contractor being a relief driver who was not familiar with the layout of the plant, itself an unmanned installation and so onsite advice was not available. The contaminated water entered the town's supplies. Its pH was as low as 3.9 and contained concentrations of aluminium and sulphate of 620 and 4500 mg/L, respectively. Within two days, many residents of Camelford complained about the poor taste of the water, skin irritation and corrosive effects of the water on plumbing and fixtures. Although the cause of the problem with the water was not solved for two days, the responsible water authority assured the town's people that the water, while tasting slightly acidic, was safe to drink. Once the cause of the problem was determined, a programme of flushing the water supplies rapidly reduced the concentration of aluminium to 1 mg/L. After several months, over 400 complaints had been made relating to skin rashes, sore throats, painful joints, memory losses and exhaustion. Following the incident, the water authority responsible was prosecuted for causing a public nuisance and fined £10 000 with additional legal costs of £25 000.

Reports were published in 1989 and 1991 following investigations by the UK government-appointed Lowermoor Incident Health Advisory Group into complaints of long-standing ill health in residents of Camelford. Both studies concluded that there was no evidence of any increase in ill health in the community related to any toxic effects of the aluminium contamination of the water. Rather, the symptoms suffered by affected individuals were largely attributed to anxiety. However, these findings are inconsistent with those of a scientific study conducted in 1991, but not published until 1999 for legal reasons, which showed that there was an association between the contamination and reduced cerebral functions in affected residents. Criticisms have been aimed at this study. The incident is still surrounded with considerable controversy, with multiple claims for damages, and a more in-depth third inquiry into possible delayed or persistent health effects was established by the UK government in 2001. This inquiry reported in 2005 and concluded that, although it was unlikely that any of the chemicals involved in the water contamination were responsible for any long-term effects. However, further studies were recommended on the effects of the contaminants on neurological health, on diseased joints in residents of the area and on the development of children below the age of 12 months at the time of the aluminium sulphate discharge. In 2013, the Lowermoor Subgroup of the Committee on Toxicity of Chemicals in Food published its final report. Its conclusions stated that the poisoning was unlikely to have caused delayed or persistent harm to health. However, later in the same year, the UK government apologised unreservedly for the unprecedented incident and the slowness of the relevant water authority in recognising the problem and taking appropriate remedial actions.

Lead

Lead is a heavy metal found as an environmental pollutant. Common sources include paints, very old water pipes, petrol fumes in some countries, industrial pollution, and some herbal/complementary 'medications' and cosmetics from the Indian subcontinent. Poisoning by lead can be acute, which is rare, or chronic, which is more common. It results from inhalation, dermal absorption or ingestion, where it is actively absorbed by the GIT through the calcium transport system (*Chapter 11*). Clinical features of lead poisoning include lethargy, abdominal discomfort, anaemia, constipation, encephalopathy and motor peripheral neuropathy.

All tissues are affected by lead poisoning, although organ pathology is associated primarily with the nervous system and the blood. In the blood, lead is concentrated in

erythrocytes, where it inhibits the activities of aminolevulinic acid (ALA) dehydratase and ferrochelatase, enzymes involved in the biosynthesis of haem (*Figure 12.27*). Inhibition of ALA dehydratase is the most sensitive measure of lead poisoning. Concentrations of lead in the blood of more than 0.49 μmol/L are associated with illness in children.

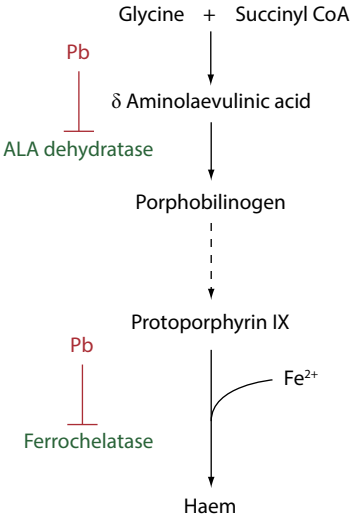


Figure 12.27 A schematic showing the enzymes of haem biosynthesis affected by lead poisoning. ALA, aminolevulinic acid.

The management of patients with lead poisoning involves identifying the lead source and, if its concentration in the blood exceeds 3.4 μmol/L, then the patients must be removed from its source. Patients are placed on chelating agents, such as ethylenediaminetetraacetic acid (EDTA; *Figure 12.28*), which is effective in both acute and chronic lead poisoning, but has to be administered intravenously. The chelating agent dimercaptosuccinic acid (DMSA) is less efficient than EDTA in removing lead but can be given orally.

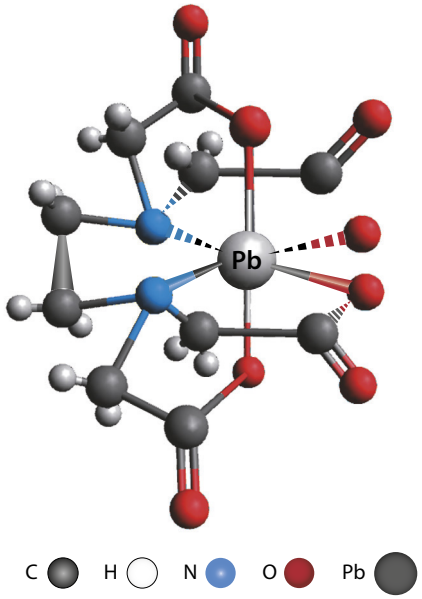


Figure 12.28 Computer generated structure of lead ethylenediaminetetraacetic acid (EDTA). The large central sphere represents the bound lead atom.

Arsenic

Chronic arsenic poisoning is associated with well water contaminated with arsenical pesticides or, classically, with murder! Arsenic occurs in several different forms, most of which are toxic, although arsenite (AsO_2^-) is much more so than arsenate ($HAsO_4^-$). Arsenite toxicity is associated with the formation of a stable complex with enzyme-bound lipoic acids (*Figure 12.29*). Indeed, arsenic poisoning can be explained by its ability to inhibit enzymes, such as pyruvate dehydrogenase, 2-oxoglutarate dehydrogenase and branched chain α-oxoacid dehydrogenase, which require lipoic acid as a coenzyme.

Chronic poisoning with arsenic is usually associated with diarrhoea, polyneuropathy and dermatitis, whereas its acute form gives rise to severe gastrointestinal pain,

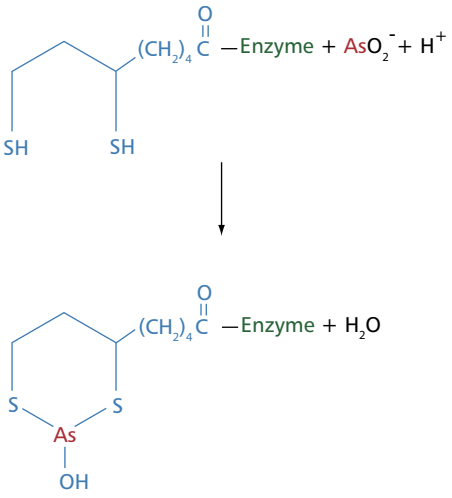


Figure 12.29 The reaction of arsenite with enzyme-bound lipoic acid.

vomiting and shock. Poisoning can be diagnosed by determining the concentration of arsenic in the hair or fingernails of the victim. Values above 0.5 μg/g in hair indicate a significant exposure to arsenic. The hair of a person chronically exposed to arsenic could have 1000 times as much as this. Treatment of arsenic poisoning is aimed at enhancing its excretion using chelating agents.

BOX 12.4 THE DEATH OF NAPOLEON: ARSENIC AND WALLPAPER

Following his defeat at the Battle of Waterloo in 1815, Napoleon Bonaparte (1769–1821) was exiled on St Helena. This is a volcanic island only about six by eight miles in size in a remote region of the south Atlantic. During most of his exile, Napoleon lived in Longwood House with a retinue of about 20. Napoleon never left St Helena and, indeed, died there in 1821 (*Figure 12.30*).

Figure 12.30 Napoleon's death mask. Courtesy of H. Ball, www.grand-illusions.com.

A postmortem of the body showed an enlarged liver and stomach lesions. It was concluded that he had died of a perforated stomach ulcer that had turned cancerous. Napoleon was initially buried on St Helena but his body was removed 20 years later and reburied at Les Invalides on the banks of the Seine in Paris, as had been his wish.

In 1952, Sten Forshufvud read an account of Napoleon's death and, given his symptoms, concluded that Napoleon may have been murdered by arsenic poisoning. White arsenic or arsenic oxide is extremely poisonous and its symptoms (stomach pains, diarrhoea, shivering and swollen limbs) can be confused with other illnesses. Indeed, a few days prior to his death, Napoleon had requested that his doctor make a full examination, particularly of his stomach, and some of his symptoms did correspond with those of arsenic poisoning.

It is possible to poison a person by slow exposure to small quantities of arsenic over an extended period. A number of Napoleon's staff had kept locks of his hair, which were subsequently passed down the generations. Hair is largely composed of keratin, a protein that contains sulphur. If arsenic is ingested, some of it will bind to the sulphur atoms and, since hair is constantly growing, it can even show how the concentrations of arsenic in the body change with time. Hair is resistant to degradation, and some of the samples of Napoleon's hair could still be analysed in the 1960s using neutron activation techniques. Several hairs were found to contain abnormal amounts of arsenic. Naturally, this led to speculation that Napoleon had been murdered. However, other possible sources of arsenic poisoning are paints and wallpapers. The pigment Scheele's green contains copper arsenite and has been used in fabrics and wallpapers since about 1770. Unfortunately, throughout the nineteenth century, many people were made ill or killed by this wallpaper. In 1893, Bartolomeo Gosio determined the pathological mechanism: if wallpapers containing Scheele's green become damp and contaminated with moulds, such as *Scopulariopsis brevicaulis*, the moulds can metabolise the arsenic salts to volatile poisonous compounds, for example trimethyl arsine, which are released in a vaporous form into the atmosphere. Following a public appeal in 1980 for help in making a radio programme, a small piece of wallpaper from the wall of the drawing room of Longwood House was located (*Figure 12.31*). The sample of paper showed a single star and its main colours were green and brown. Gold and green were the imperial colours, but it is possible that the brown was originally gold in colour. More significantly, X-ray fluorescence spectroscopy analysis showed

the wallpaper to contain appreciable quantities of arsenic. Longwood House was notoriously damp, and the paper would have degraded to release volatile arsenic compounds. Further, many of its inhabitants other than Napoleon had apparently also become ill and complained of the 'bad air', so it is possible that he might have been a victim of British wallpaper makers rather than a deliberate murder plot. However, the amount of arsenic released could not have been large, and apparently was insufficient to have killed Napoleon, although once he did become ill with a stomach ulcer, the arsenic could have exacerbated his condition.

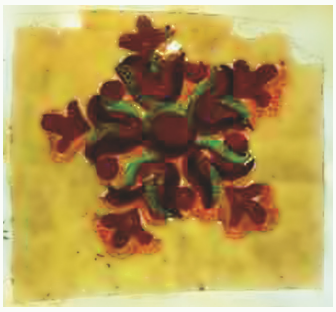


Figure 12.31 A sample of Napoleon's wallpaper. Courtesy of H. Ball, www.grand-illusions.com.

Cadmium

Chronic cadmium toxicity may occur in workers exposed to fumes in cadmium-related industries, although concentrations of cadmium are twice as high in tobacco smokers as in nonsmokers. Concentrations in the serum greater than 90 nmol/L are associated with toxicity. Cadmium poisoning causes influenza-like symptoms, such as chills, fever, muscular aches, nausea, vomiting, abdominal pain and diarrhoea. However, these symptoms may resolve after a week, provided there is no respiratory damage. More severe exposures can cause bronchitis and pulmonary oedema and occasionally cardiovascular collapse. Long-term exposure may lead to nephrotoxicity with proteinuria, bone disease and hepatotoxicity. The treatment of cadmium poisoning usually involves removal from the source of exposure. Treatment with chelating agents is not effective because the soluble form of cadmium damages the kidneys.

Mercury

Mercury poisoning can be acute or chronic. It occurs following exposure to mercury vapour, inorganic salts or organic compounds of mercury. Mercury poisoning is primarily occupational but can be caused by contaminated food (*Margin Note 12.6*). The clinical features of acute mercury poisoning include coughing, bronchiolitis, pulmonary oedema, pneumonitis, peripheral neuropathy and neuropsychiatric problems. Chronic mercury poisoning causes anorexia, insomnia, impaired memory, paraesthesiae of the lips and extremities, renal tubular damage and sweating. Mercury poisoning is usually diagnosed by determining the concentration of mercury in serum and urine. Urine mercury/creatinine ratios are often used to assess exposure. Ratios of 40–100 nmol/mmol require monitoring and further investigation, while values greater than 100 nmol/mmol require the patient to be removed from the source of mercury. Acute mercury poisoning is treated with chelating agents such as dimercaprol to increase its excretion into bile and urine.

Margin Note 12.6
Mercury contaminated food

Methylmercury is an organic form of mercury that has been used to preserve seed grain. Methylmercury poisoning can be caused by eating meat from animals that have been fed treated seedgrain. An outbreak of mercury poisoning began in the 1950s in Minamata Bay in Japan and affected over 3000 villagers who ate fish contaminated with methylmercury.

12.6 CASE STUDIES

CASE STUDY 12.1

Helen, an 18 year old, was admitted to the accident and emergency department of her local hospital. She had been found in her bedroom with an empty bottle of aspirin tablets. On admission, Helen was in a confused state, sweating and breathing heavily. There was clinical evidence for mild dehydration. She had a pulse rate of 112/min and blood pressure of 110/70 mmHg (14.7/9.3 kPa). Her body temperature was 39°C. Laboratory investigations give the following profile (reference ranges are shown in parentheses):

Na^+	132	(135–145 mmol/L)
K^+	3.4	(3.6–5.0 mmol/L)
HCO_3^-	11	(22–30 mmol/L)
Urea	11	(3.3–6.7 mmol/L)
Glucose	3.5	(2.8–6.0 mmol/L)
Salicylate	4.6	(up to 2.5 mmol/L is therapeutic)
pH	7.20	(7.36–7.44)
PCO_2	3.6	(4.5–6.0 kPa)
Prothrombin time	18	(14 seconds)

Questions

(a) What can be concluded from these clinical data?
(b) What treatment would be suitable?

CASE STUDY 12.2

Olga, a 25 year old, was admitted to hospital after being found unconscious by her flatmate. It was reported that empty vodka and paracetamol bottles had also been discovered in her flat. On admission to an accident and emergency department, Olga's breath smelt of alcohol and her serum paracetamol concentration was 105 mg/mL. Her liver function tests were normal, as was her prothrombin time. Serum alanine aminotransferase activity showed a transient increase over the next few days but gradually fell back within the reference range.

Questions

(a) Comment on these clinical findings.
(b) What treatment should Olga be given?

CASE STUDY 12.3

Ed, a 58-year-old building site worker, has been a heavy alcohol drinker for over 30 years. He was admitted to hospital after collapsing outside a public house. He smelt strongly of alcohol and was confused and unsteady on standing. He complained of fatigue, nausea (particularly early in the morning) and a loss of appetite. A physical examination detected abdominal tenderness and enlargement of the liver and gastrointestinal varicose veins. His blood alcohol concentration was 79 mmol/L. Ed's Ca^{2+} and Mg^{2+} blood concentrations were low and his clotting time extended. His urinary urea excretion was also subnormal.

Maureen, a 52 year old, was admitted into hospital at a similar time. She presented with general GIT problems and frequent diarrhoea. She was subjected to considerable stress at work and admitted she had been drinking 'rather heavily'.

Liver function tests were performed on sera from both patients. The results are shown below (reference ranges are shown in parentheses):

	Ed	Maureen	
Total protein	67	78	(60–84 g/L)
Albumin	31	39	(35–50 g/L)
Total bilirubin	57	14	(3–15)
Alkaline phosphatase	720	335	(100–300 U/L)
Alanine transaminase	34	91	(5–35 U/L)
Aspartate transaminase	41	162	(10–40 U/L)
γ-Glutamyltransferase	780	455	(7–45 U/L)

Questions

(a) Suggest why the tests above are particularly indicative of the state of health of the liver.
(b) Does Ed or Maureen show the greater liver degeneration?

CASE STUDY 12.4

John's girlfriend found him in his closed garage unconscious in his car with the engine running. He was taken to the hospital immediately. The following blood results were obtained (reference ranges are shown in parentheses):

H^+	50	(36–46 nmol/L)
HCO_3^-	15	(22–30 mmol/L)
PCO_2	3.5	(4.5–6.0 kPa)
Haemoglobin	150	(130–180 g/L)
Lactate	10	(0.65–2.0 mmol/L)

Questions

(a) Suggest a plausible explanation for these clinical data.
(b) How should John be treated?

CASE STUDY 12.5

Abdul, a 61-year-old man with a history of type 2 diabetes, was admitted to hospital suffering from abdominal pain, weakness and reduced appetite, which he had developed over the last four months. On physical examination, he did not have jaundice but did have pallor. He reported no history of vomiting. The following blood results were obtained (reference ranges are shown in parentheses):

Haemoglobin	105	(140–180 g/L)
Erythrocyte count	3.33×10^{12}	($4.4–5.9 \times 10^{12}$/L [men])

A blood smear showed microcytic hypochromic anaemia with basophilic stippling.

On further questioning, Abdul admitted to taking herbal medication for the treatment of his diabetes over the last 12–15 months. His blood lead was measured and gave the following result (normal values are shown in parentheses):

Blood lead	3.2	(<0.48 μmol/L)

Questions

(a) What could be the possible explanation for these findings?
(b) How should Abdul be treated?

12.7 SUMMARY

The body is exposed constantly to many chemicals, some of which have the potential to harm. These chemicals may be ingested with food and drink, may be therapeutic or recreational drugs or may be toxins from microorganisms or the abiotic environment. Xenobiotics may be detoxified in the liver, by conversion to more water-soluble compounds, which are then excreted through the kidneys. Sometimes this process results in more harmful compounds being produced, as is the case with paracetamol poisoning. Clinical investigations of suspected poisoning include liver function tests, analysis of blood gases and determination of acid–base status.

The most common cases of poisoning involve accidental or deliberate overdoses with common analgesics such as paracetamol or aspirin, but abuse of alcohol is an increasing problem in developed countries, leading to liver damage in the patient. In addition, carbon monoxide poisoning may be insidious and fatal in badly ventilated houses warmed by faulty gas heaters. Treatment of poisoning first involves identification and removal of the poison; thereafter, treatment depends on the nature of the poison involved. Unfortunately, some poisons, such as paraquat, are invariably fatal.

QUESTIONS

1. Which of the following poisonings produces a combination of respiratory alkalosis and metabolic acidosis?
(a) Paracetamol.
(b) Ethylene glycol.
(c) Aspirin.
(d) Lead.
(e) Carbon monoxide.

2. Which of the following is the toxin produced during paracetamol poisoning?
(a) Paracetamol-3-mercapturic acid.
(b) Glutathione.
(c) *N*-Acetylbenzoquinoneimine.
(d) UDP-Glucuronyltransferase.
(e) Trimethylarsine.

3. What are the possible prognoses for a patient who has taken an overdose of paracetamol if treated within hours or if treatment is delayed for several days?

4. Disulfiram is an inhibitor of ALDH. It was used as the drug Antabuse to deter alcohol consumption in chronic alcoholics. Suggest a plausible reason why treatment with it should deter alcoholics from consuming alcoholic drinks.

5. Suggest why alcohol should be consumed with caution or avoided when taking other drugs.

6. A 10-year-old Asian girl was admitted to hospital where she presented with vomiting, nausea and neurological symptoms. She had been using a traditional cosmetic called surma that was applied around the eyes. What laboratory investigations should be applied to this patient?

7. A 50-year-old scientific instrument technician repaired old scientific gauges that often contained mercury. One winter, he complained of weight loss, weakness, symptoms that resembled influenza and depression. On examination, he had muscle weakness and wasting. A urine sample contained 180 nmol Hg/mmol creatinine. How should this patient be treated?

8. Which of the following statements are TRUE?
(a) 2,4-Dinitrophenol breaks the link between glycolysis and the electron transport chain, which produces ATP.
(b) Drug receptors are invariably proteins.
(c) The sinusoids of liver lobules facilitate direct contact between blood and hepatocytes.
(d) Mixed function oxidases require NAD^+ as a coenzyme and dioxygen for activity.
(e) Glutathione is able to detoxify paracetamol by binding to it, forming NAPQI.
(f) A large overdose of aspirin can result in the excretion of salicylate in urine.
(g) Acidic provisional urine produced in the kidneys promotes the ionisation of salicylic acid, allowing it to be reabsorbed from the tubules.

(h) In a healthy individual, the major route for detoxifying ethanol is by P-450-catalysed oxidation.
(i) The toxic effects of ethylene glycol largely arise following its oxidation to oxalate.
(j) Carbon monoxide binds to haemoglobin about 240 times more strongly than dioxygen.

9. Which of the following are inappropriate pairings?

(a)	Papaverine	Mild–moderate pain
(b)	Esomeprazole	Gastric reflux
(c)	Cytochrome P-450	$FADH_2$
(d)	Glycogen phosphorylase	Phosphorylation of glycogen
(e)	Stage III paracetamol poisoning	Renal and hepatic failure
(f)	Ethanal	Ethanol dehydrogenase
(g)	Phenobarbitone	Produces longer-lasting effect than most barbiturates
(h)	Alveolar epithelial cells	Active accumulation of paraquat
(i)	Aluminium	Osteodystrophy and encephalopathy
(j)	Lead	Aminolevulinic acid dehydratase and ferrochelatase

FURTHER READING

Alasfar RH and Isaifan RJ (2021) Aluminium environmental pollution: the silent killer. *Environ Sci Pollut Res Int* **28**:44587–44597.

Barlow NL and Bradberry SM (2023) Investigation and monitoring of heavy metal poisoning. *J Clin Pathol* **76**:82–97.

Desborough MJR and Keeling DM (2017) The aspirin story – from willow to wonder drug. *Br J Haematol* **177**:674–683.

Fisher ES and Curry SC (2019) Evaluation and treatment of acetaminophen toxicity. *Adv Pharmacol* **85**:263–272.

Gawarammana IB and Buckley NA (2011) Medical management of paraquat ingestion. *Br J Clin Pharmacol* **72**:745–757.

Jones AW (2021) Clinical and forensic toxicology of methanol. *Forensic Sci Rev* **33**:117–143.

Mitra P *et al*. (2017) Clinical and molecular aspects of lead toxicity: an update. *Crit Rev Clin Lab Sci* **54**:506–528.

Nurchi VM *et al*. (2020) Arsenic toxicity: molecular targets and therapeutic agents. *Biomolecules* **10**:235. doi: 10.3390/biom10020235.

Ramachandran A and Jaeschke H (2018) Acetaminophen toxicity: novel insights into mechanisms and future perspectives. *Gene Expr* **18**:19–30.

Ross JA *et al*. (2022) Toxic alcohol poisoning. *Emerg Med Clin North Am* **40**:327–341.

Stucki D and Stahl W (2020) Carbon monoxide – beyond toxicity? *Toxicol Lett* **333**:251–260.

Yang L *et al*. (2020) Toxicity of mercury: molecular evidence. *Chemosphere* **245**:125586. doi: 10.1016/j.chemosphere.2019.125586.

Disorders of the blood

13

OBJECTIVES

After studying this chapter you should be able to:

- list the components of blood
- describe the structures of blood cells, erythrocytes and platelets
- discuss the functions of haemoglobin
- explain how blood clotting occurs
- list examples of haemoglobinopathies and anaemias
- explain the consequences of the mutations underlying sickle cell anaemia and thalassaemias
- describe how some types of anaemias are diagnosed and treated
- describe some clinical disorders associated with clotting factors

13.1 INTRODUCTION

Blood is an electrolyte- and protein-rich fluid called **plasma** in which erythrocytes, leukocytes (sometimes called red and white blood cells, respectively) and platelets are suspended (*Figure 13.1*).

Erythrocytes, leukocytes and platelets are sometimes referred to as **formed elements**. They constitute about 40–45% of blood volume. Blood is pumped by the heart around the body through the arteries that supply the capillaries and is returned to the heart in the veins (*Chapter 14*). The main functions of the blood are to distribute oxygen, nutrients, hormones and other signalling molecules between tissues and to remove carbon dioxide and other waste products. Plasma contains the proteins of both clotting and immune systems (*Chapters 4* and *5*). It also contains a range of other plasma proteins, additional to the clotting and immune system proteins, and nutrients, such as glucose, waste materials (for example urea) and a number of electrolytes. If it is allowed to clot, the clear straw-coloured liquid remaining after removal of the clot is called **serum**. The volume of blood in males and females is approximately 78 and 66 mL/kg body weight, respectively. The compositions of the blood and plasma are given in *Table 13.1*.

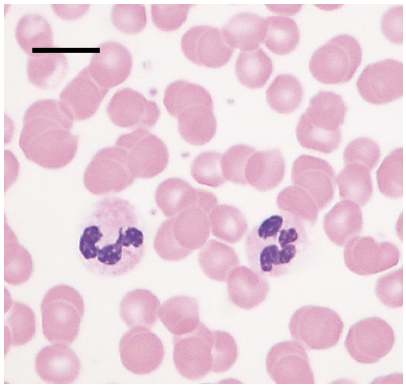


Figure 13.1 Photomicrograph of a blood smear. The biconcave shape of erythrocytes makes their centres appear lighter in colour. Erythrocytes predominate, with only two leukocytes visible. ⏎

Plasma	**Concentration (g/L)**
Total protein	60–83
Fibrinogen	1.9–8.0
Albumin	34–54
Globulins (excluding fibrinogen)	19–35
Erythrocytes, leukocytes and platelets	**Number/L**
Erythrocytes	
Male	$4.4–5.9\times10^{12}$
Female	$3.8–5.2\times10^{12}$
Leukocytes	$4–11\times10^{9}$
Platelets	$2.5–5.0\times10^{9}$

Table 13.1 Composition of the blood. ⏎

DOI: 10.1201/9781003296836-13

In a text of this size it is not, of course, possible to discuss each type of blood disorder, and attention will focus only on the major types of diseases likely to be normally encountered.

13.2 ERYTHROCYTES, LEUKOCYTES AND PLATELETS

All of the formed elements of blood originate from pluripotent stem cells in bone marrow (*Figure 13.2*). Chemical signals, such as cytokines (*Chapter 4*), direct primordial stem cells to develop in different ways to produce:

- erythrocytes
- leukocytes (*Chapter 4*)
- megakaryocytes, the precursors of platelets

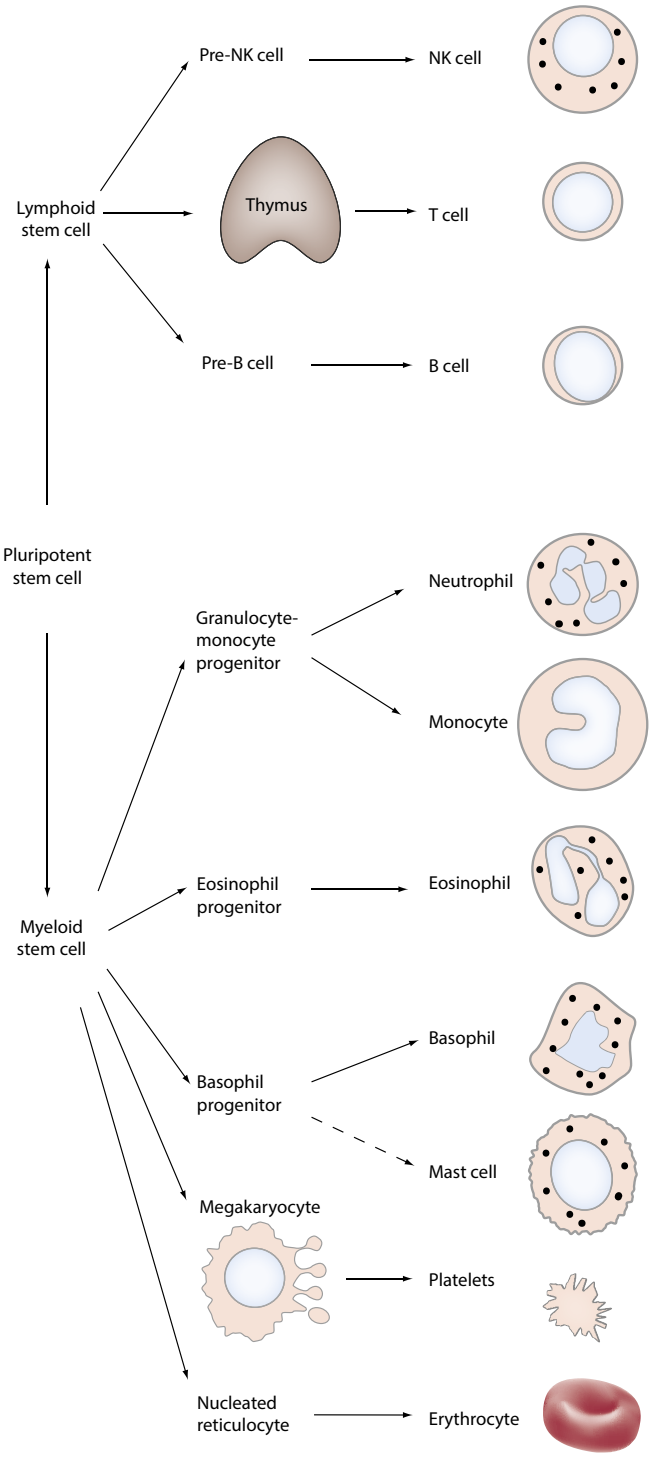


Figure 13.2 Erythrocytes, leukocytes and platelets all originate in the bone marrow where progenitor stem cells divide, and their daughter cells differentiate to produce the different formed elements. This differentiation is controlled by growth factors. See *Chapter 4* for more information on lymphocytes (natural killer (NK), T and B cells).

Normoblasts are erythroid cells that arise from divisions of pluripotent stem cells. Eventually these lose their nuclei giving rise to **reticulocytes** that contain mRNAs for globins and are still able to synthesise haemoglobin (Hb), and which are the precursors of the erythrocytes. Reticulocytes circulate in the blood for one to two days before maturing to erythrocytes and normally constitute 1–2% of the circulating 'red' types of formed elements. Erythrocytes are the most numerous 'cells' in the blood, with approximately 2×10^{11} formed daily.

Erythrocytes

Erythrocytes lack nuclei and other organelles and are biconcave in shape (*Figure 13.3*), which increases their surface area/volume compared with a simple sphere, so increasing the area for O_2 absorption. They have diameters of about 7 µm, the same as capillaries. Hence, as they pass through capillaries, they become deformed, but they regain their shape in larger blood vessels (*Chapter 14*). This plasticity of erythrocytes depends to a large extent upon their specialised flexible cytoskeleton. Mutations in the genes encoding the proteins of their cytoskeleton lead to blood disorders, such as elliptocytosis and spherocytosis (*Chapter 16*; see also *Figure 16.38*). Erythrocytes circulate for about 120 days and are then removed from circulation and destroyed by macrophages in the liver and spleen.

Haemoglobin, the red-coloured protein found in the erythrocytes, carries dioxygen (O_2) and also plays a significant role in the buffering that maintains the pH of plasma at 7.4±0.1 (*Chapter 9*). The iron-containing haem is removed from the Hb of defunct erythrocytes and its porphyrin ring is converted to bilirubin, which is excreted in the bile (*Chapter 11*). The iron is conserved and recycled. Iron circulates in the blood bound to its transport protein transferrin (see *Figure 11.34*) and is stored in liver and spleen cells in ferritin (*Figure 13.4*). An excess breakdown of Hb produces more bilirubin than can be excreted and this accumulates in the tissues causing jaundice (*Box 13.1* and *Chapter 11*).

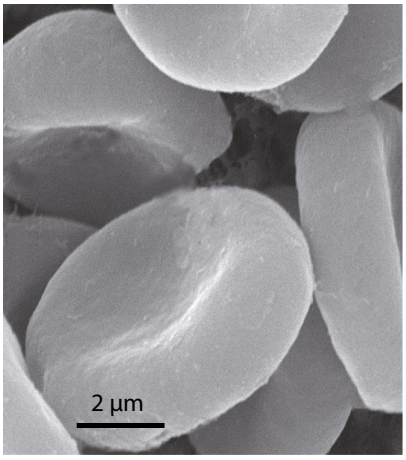


Figure 13.3 Scanning electron micrograph of erythrocytes. Courtesy of the Centers for Disease Control and Prevention/Dr J.H. Carr.

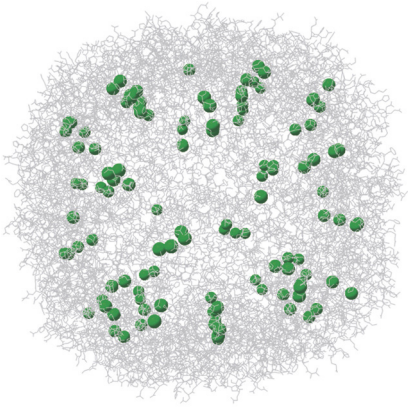


Figure 13.4 Molecular model of ferritin. The stored Fe complexes are represented as green spheres. When filled, one ferritin molecule can store more than 2000 iron atoms as hydrated ferric hydroxide. PDB file 4OYN.

BOX 13.1 IRON METABOLISM

Only about 10% of the average daily intake of iron, approximately 20 mg in the UK, is absorbed, mostly by the duodenum and jejunum (*Chapter 11*), although more is absorbed in pregnancy and in iron deficiency anaemia (*Section 13.6*). There are obviously increased demands for iron in growth periods and menstruation, when about 0.7 mg of iron is lost daily, and in pregnancy (*Chapter 7*). Haem iron from haemoglobin (Hb) and myoglobin in red meats is better absorbed than nonhaem iron. Absorption is controlled by the mucosal cells of the small intestine and the iron is to some extent stored in these cells before being passed to hepatic portal blood bound to transferrin (see *Figure 11.34*). Stored iron may be lost when these cells are shed. Transferrin can carry two atoms of iron per molecule but is only about one-third saturated on average. The iron is detached from the transferrin in the bone marrow when the protein interacts with specific receptors on erythroblasts and reticulocytes supplying their iron needs for Hb synthesis. Although 60–70% of the body's iron is found in the circulation as Hb, with more in cytochromes and other iron proteins, some is stored in the protein ferritin (*Figure 13.4*). This protein is found in most cells, but particularly in those of the liver and spleen.

The body does not have an excretory route for iron, but this is not normally a problem unless repeated blood transfusions (*Chapter 6*) are given, for example to patients with sickle cell anaemia (SCA, also called sickle cell disease [SCD]) or a thalassaemia (*Section 13.7*), because as the transfused erythrocytes are broken down the iron may form haemosiderin. This consists largely of iron oxide and accumulates mainly in the liver and spleen but also in other tissues. The genetic condition haemochromatosis can also result in the deposition of haemosiderin (*Figure 13.5*).

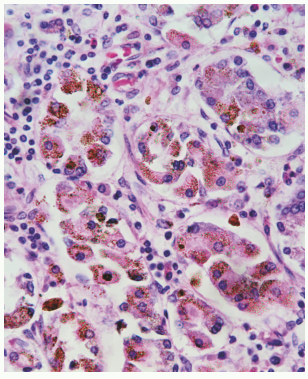


Figure 13.5 Photomicrograph showing the accumulation of haemosiderin in stomach cells of a patient with haemochromatosis. Courtesy of Shutterstock.com/Jubal Harshaw.

Leukocytes

There are several types of leukocytes in the blood, each performing specific function(s). The total number of leukocytes in an adult is given in *Table 13.1*, but there is considerable variation in numbers (*Box 13.2*).

BOX 13.2 COUNTING CELLS

In hospital laboratories, cells are counted using sophisticated automatic instruments, which can simultaneously give numerous data of clinical relevance (*Figure 13.6*). However, in teaching laboratory exercises for example, blood cells can be counted using a haemocytometer (*Figure 13.7*).

The blood sample is diluted and then pipetted into a glass chamber of known volume. The number of cells within the indicated grid is counted under the microscope, and this number multiplied by the dilution gives the number of cells in the sample. Different dilutions need to be done for erythrocytes and leukocytes, given the differences in their numbers.

Automated cell counters and flow cytometers can also be used to count eukaryotic cells grown in culture (*Chapter 6*).

Figure 13.6 A blood analysis system that can determine all haematological indices automatically. Courtesy of L. Jones and Drs C. Chaloner and E. Cowdell, Manchester University NHS Foundation Trust, UK.

Leukocytes were originally classified on the bases of their size, which microscopic stains they took up, whether they had a granular cytoplasm and the shapes of their nuclei (see *Chapter 4*).

Types of leukocytes

The three main types of leukocyte are:

- **polymorphonucleocytes** (PMNs, sometimes referred to as *polymorphs*)
- **lymphocytes**
- **monocytes**

Polymorphonucleocytes are further subdivided into **neutrophils** (57% of the total leukocyte population) that contain neutral staining granules, **eosinophils** (3.5%) that contain acid-staining granules, and **basophils** (0.5%) with basic staining granules (*Figure 13.2*, see also *Figure 4.15*). Polymorphonucleocytes can secrete cytokines, a number of which are mediators of inflammation as described in *Chapter 4*. Neutrophils can migrate to areas of infection and phagocytose bacteria. Eosinophils seem to be more concerned with dealing with parasites, but their number also increases in allergic diseases. Basophils, and the similar mast cells, which are found mostly in the skin, can release the histamine stored in their granules, and this also contributes to some types of allergic responses (*Chapter 5*).

Lymphocytes are key components of the immune system. Monocytes circulate in the blood for about 72 hours before they enter solid tissues and differentiate into macrophages. Both lymphocytes and monocytes play key roles in immune defence and inflammation (*Chapter 4*).

Platelets

Platelets are vesicle-like structures about 3 μm in diameter with a volume of 7 fL (femtolitres). They are formed by the fragmentation of large precursor cells called megakaryocytes and, like erythrocytes, are 'incomplete cells', although they contain numerous granules, some organelles and a tubular system (*Figure 13.8*). Their lifespan is 10–12 days and they function in blood clotting or haemostasis (*Section 13.4*).

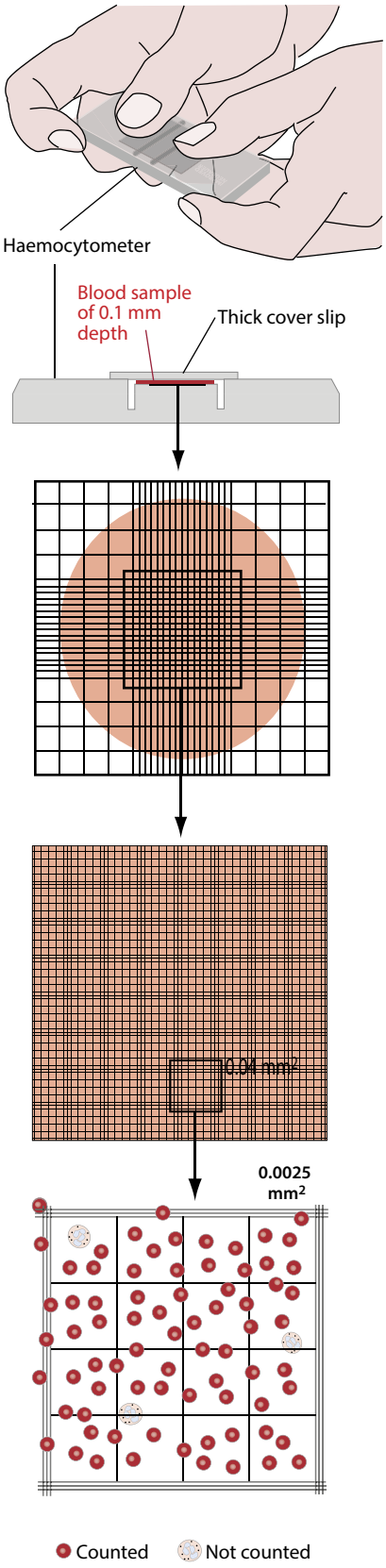


Figure 13.7 The use of a haemocytometer to determine the concentrations of erythrocytes and cells in blood.

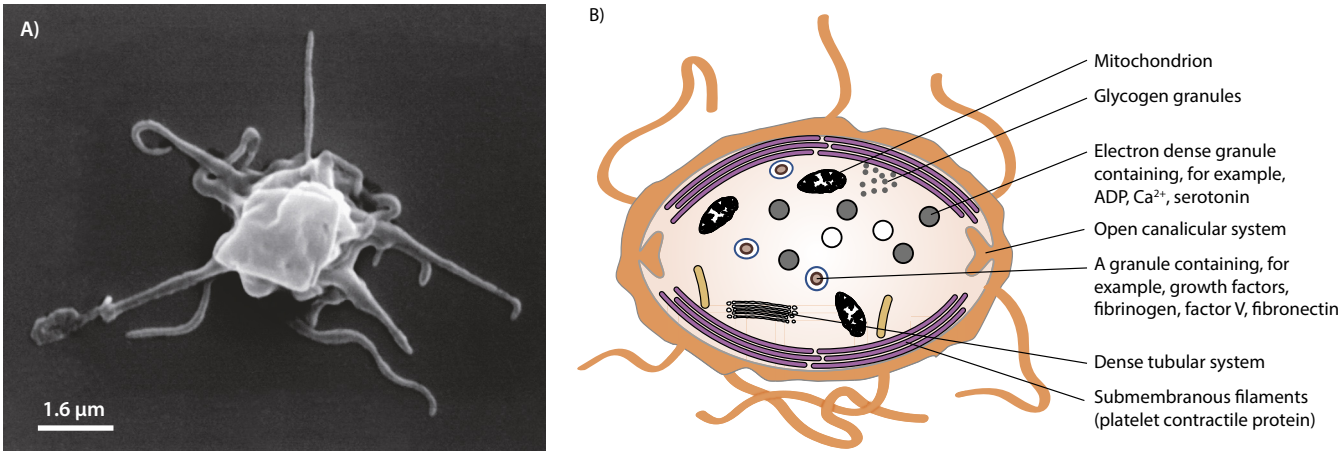


Figure 13.8 (A) Electron micrograph of an activated platelet. Courtesy of Wellcome Collection/David Gregory and Debbie Marshall. (B) Schematic of a blood platelet. The platelet is packed with granules that have a major role in blood clotting.

13.3 HAEMOGLOBINS

Haemoglobin was one of the first proteins to have its structure determined. Indeed, because Hb is so significant medically, it is fair to say that more is known about it than any other protein. Its functioning is vital because the relatively low solubility of oxygen in water means that insufficient amounts can be carried in solution by the blood. The binding of O_2 to Hb in erythrocytes increases the oxygen-carrying capacity several thousand-fold. Furthermore, the body has a limited capacity to work anaerobically and so O_2 is essential to support metabolism. Even a small reduction in the amount of Hb in the blood leads to **anaemia**, which causes serious clinical problems (*Section 13.6*). Anaemia can result from there being too few erythrocytes or from each erythrocyte having too little or a defective Hb. These conditions may be a consequence of a lack of iron. In all cases, clinical intervention is usually necessary.

Haemoglobin's M_r is about 64 000. It consists of four subunits, each containing an iron-containing haem group (*Figure 13.9*). Each Hb molecule can carry up to four O_2 molecules.

Margin Note 13.1
Platelet concentrates

Platelet concentrates (see *Figure 6.2*) can be prepared by centrifugation and may be stored for up to five days. Such concentrates are used to treat patients suffering from, for example, thrombocytopenia (*Section 13.9*), who have insufficient platelets, and to prevent bleeding in patients with bone marrow failure.

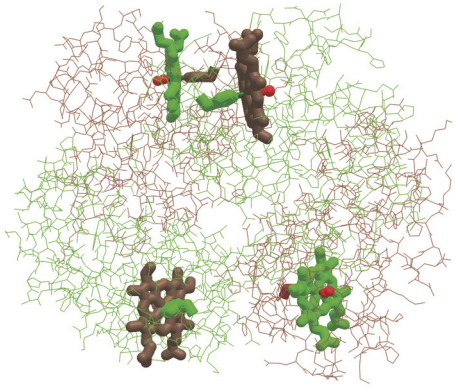


Figure 13.9 Molecular model of adult oxyhaemoglobin. The α and β subunits are shown in green and brown, respectively, with their haem groups shown in contrasting colours. The bound dioxygens are represented by red spheres. Note how the four subunits interact closely with each other. PDB file 1GZX.

Oxygen is taken up as the blood passes through the lungs and is transported to all parts of the body allowing respiration (the oxidation of fuels) to occur in mitochondria (*Chapter 16*). The iron in the haem group of Hb remains in the ferrous (Fe^{2+}) state throughout. Should the iron become oxidised to the ferric condition (Fe^{3+}), **methaemoglobin** is formed, which is incapable of carrying O_2 because a bound H_2O blocks its binding site. This oxidation happens to a small extent continuously, so that normal blood always contains about 1% methaemoglobin. However, NADH-cytochrome b_5 reductase (methaemoglobin reductase; *Figure 13.10*), present in erythrocytes, constantly catalyses its reduction back to Fe^{2+}. The unfortunate individuals with genetic deficiencies of this enzyme suffer from methaemoglobinaemia, which can cause severe clinical problems (*Section 13.5*).

Figure 13.10 Molecular model of the erythrocyte NADH-cytochrome b_5 reductase (methaemoglobin reductase) shown in green with its bound flavin adenine dinucleotide (FAD) coenzyme depicted in red. PDB file 1UMK.

Key properties of haemoglobin

Haemoglobin molecules are roughly spherical, with the four subunits (see *Figure 13.9*) fitting tightly together. The subunits of Hb are all similar but occur in two identical pairs.

Thus, for example, adult Hb, referred to as HbA, can be described as $\alpha_2\beta_2$. Embryonic and fetal Hbs also occur. The haem groups are hydrophobic and sit inside hydrophobic clefts in the protein. The iron atom within each haem binds one O_2 molecule, meaning Hb binds four O_2 successively. The strength of this binding, that is, the oxygen affinity, can be measured. In the lungs, where there is abundant O_2, Hb can bind up to its maximum of four and become O_2 saturated. In other tissues, the O_2, however, must 'unbind' and be released. The strength of binding is therefore critical: too weak and Hb would be ineffective as an O_2 carrier, too tight and the tissues would not be supplied because it would not be released from the oxyhaemoglobin.

The binding and release of O_2 is not straightforward. The capture of an O_2 by one subunit induces a small change in the shape of the protein, making the binding of the next O_2 to another subunit slightly easier, that is, the strength of binding changes with each successive addition. The consequence is that the graph of bound oxygen against oxygen concentration (the oxygen binding curve) is S or sigmoidal in shape (*Figure 13.11*). Hence, in the lungs, the HbA molecule can become nearly 100% saturated with oxygen while, in other tissues, it releases almost all of it.

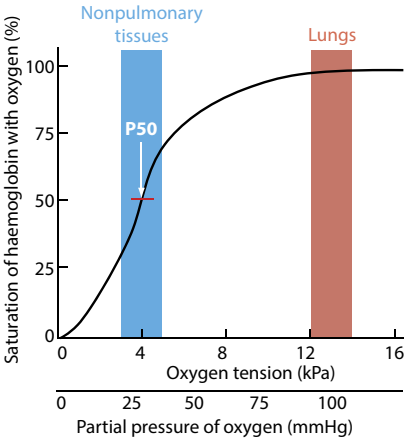


Figure 13.11 The oxygen binding curve of haemoglobin; note its sigmoidal shape, indicating that the affinity for O_2 changes as each successive O_2 binds. Thus, in the lungs, where the oxygen tension is highest, the haemoglobin becomes almost saturated with O_2. In other tissues, where the oxygen tension is lower, the haemoglobin can release almost all of its O_2.

Embryonic and fetal haemoglobins

The embryonic and fetal forms of Hb differ slightly from HbA. There are several types of embryonic Hbs present early in embryonic life but at about six weeks there is a switch to fetal Hb (HbF). Fetal Hb has two α subunits, as in the adult, but two γ subunits ($\alpha_2\gamma_2$), and there are two different types of γ subunit. Embryos and fetuses obtain their oxygen from the maternal blood in the placenta. Thus, their Hb must become saturated with O_2 at lower tensions than the maternal HbA so they can obtain it from the mother (*Figure 13.12*). This is possible because embryonic and fetal Hbs have greater affinities for O_2 than HbA (*Figure 13.13*). Adult Hb production starts in the fetus shortly before birth and by 30 weeks of age it would normally have replaced all the HbF. The change in production from HbA to HbF involves the protein BCL11A (*Figure 13.14*), a transcription factor that recognises and binds to the γ globulin gene promoter. This switches off production of γ globulin, that is, HbF, allowing it to be replaced by HbA.

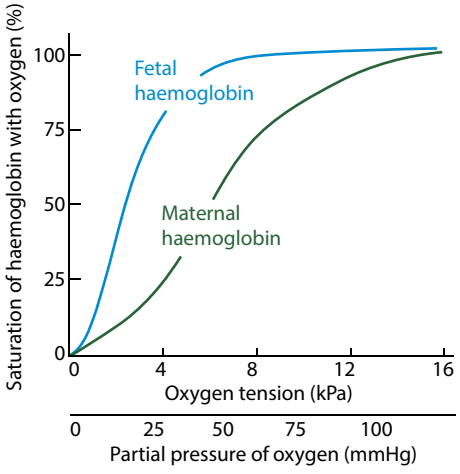


Figure 13.12 Graph demonstrating the higher affinity of fetal haemoglobin for O_2 compared with adult (maternal) haemoglobin. This feature allows the fetus to obtain O_2 from the maternal blood.

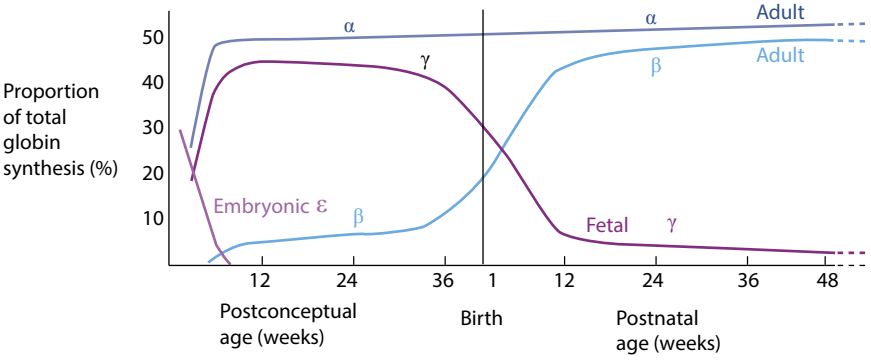


Figure 13.13 The production of human globins during development. There are several types of β globins in the embryo and fetus. Any given erythrocyte contains only one type of α and one type of β globin.

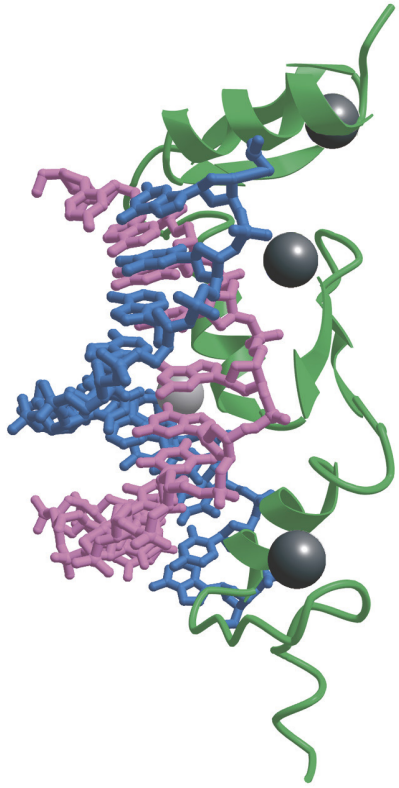


Figure 13.14 Molecular model of BCL11A (green) binding to a portion of DNA (blue and magenta). The Zn^{2+}- and Mg^{2+}-forming parts of the protein structure are shown in dark and silver grey, respectively. PDB file 6U9Q.

13.4 HAEMOSTASIS AND BLOOD CLOTTING

The circulatory system is self-sealing. Haemostasis rapidly stops all but the most catastrophic bleeding in normal individuals. If the lining of a blood vessel is damaged, eventually a platelet plug is formed that prevents further blood loss. Blood clotting is then initiated, forming a clot that seals the wound, and healing can then begin.

Clotting factors

Clotting involves a group of plasma proteins, usually called clotting factors but sometimes coagulation factors. They act in sequence to activate the next in line, and the result is the conversion of soluble fibrinogen to the insoluble fibrin. The polymerisation of the fibrin forms a clot at the site of the damage but, critically, not elsewhere (*Figure 13.15*).

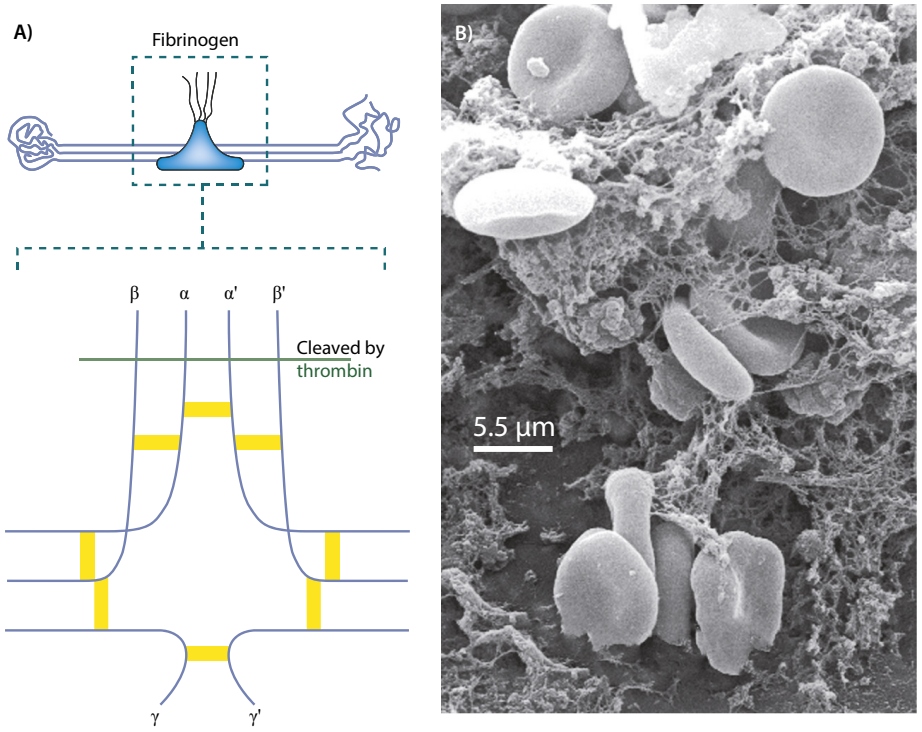


Figure 13.15 (A) Schematic showing the structure of fibrinogen and its conversion to fibrin. Fibrinogen has M_r about 340 000 and consists of two units, each containing three polypeptides (α, β and γ) joined together at their amino termini by a number of disulphide bonds. Short lengths of the amino terminal regions of the two α and two β polypeptides project outwards, and cleavage by thrombin, as indicated, allows the resulting fibrin molecules to aggregate to form a 'soft clot'. This is subsequently strengthened by the cross-linking action of transglutaminase, factor XIIIa. (B) Electron micrograph of erythrocytes embedded in fibrin. Courtesy of the Centers for Disease Control and Prevention/Dr J.H. Carr.

About 20 clotting factors are produced, mainly by the liver, and circulate in the plasma as inactive proteins (*Table 13.2*). They were called factors long before anything was known of their chemical nature and are identified by Roman numerals, although some also have common names. However, in general, factors I–IV are referred to by their common names, while the Roman numerals are used when discussing factors V–XIII. Unfortunately, as described below, the numbering of the factors does not follow any logical sequence with reference to the clotting processes. The activated forms of factors are distinguished by a lower case 'a'; hence, for example, the activation of factor VIII produces VIIIa. Many factors become proteases when activated but, as each factor catalytically activates the next one in the clotting sequence, a small amount of the first factor produces a very large reaction at the end of the pathway. Thus, the cascade of clotting reactions is an amplification of an initial small stimulus and leads to the production of large amounts of fibrin, forming a clot.

Margin Note 13.2 Hereditary persistence of fetal haemoglobin

In *hereditary persistence of fetal haemoglobin* (HPFH), the HbF is not replaced with HbA. This is presumably due to a failure of the switching mechanism involving BCL11A that normally occurs around the time of birth. Although individuals with HPFH have high concentrations of HbF, the condition does not cause any major haematological abnormalities and does not prevent those affected from having a normal life.

Margin Note 13.3 Biochemical activation

Compare the amplification mechanism shown in blood clotting with that displayed in complement activation (*Chapter 4*) and hormone stimulation (*Chapter 7*).

Factor number	Name	Active form/function	Associated diseases/clinical conditions
I	Fibrinogen	Fibrin subunit	Afibrinogenaemia (uncommon)
II	Prothrombin	Serine protease	Defective function (extremely rare)
III[a]	Tissue factor (formerly thromboplastin)	Receptor	–
IV	Ca^{2+}	Essential cofactor for many of the clotting factors	–
V	Proaccerin/labile factor	Cofactor role	–
VI	Initially wrongly identified, now known to be Va	–	–
VII	Proconvertin	Serine protease	Deficiency known (but rare)
VIII	Antihaemophilia factor	Cofactor	Haemophilia A (classical haemophilia; see *Section 13.9*)
IX	Christmas factor	Serine protease	Christmas disease (haemophilia B)
X	Stuart–Prower/Stuart factor	Serine protease	Deficiency known (but extremely rare)
XI	Plasma thromboplastin antecedent (PTA)	Serine protease	Rare, primarily found in Ashkenazi Jews
XII	Hageman factor	Serine protease	Some tendency to abnormal bleeding; may result in a form of angioedema
XIII	Fibrin stabilising factor	Transglutaminase, a serine endopeptidase	Rare, poor wound healing, scarring
XIV	Prekallikerin (F, Fletcher)	Zymogen precursor of kallikrein	Deficiency associated with an increased activated partial thromboplastin time (aPTT) but not abnormal bleeding
XV[b]	HMWK	Cofactor role	–
XVI	von Willebrand factor (vWF)	Binds to VIII mediating platelet adhesion	vWB disease (up to 125 in 10^6)
XVII	Antithrombin III	Inhibitor of thrombin and other serine protease	Mutations in *AT3D* leading to defective or deficient antithrombin III results in thrombophilia (recurrent venous thromboses) and thromboembolism (blood vessels blocked by abnormally deposited clots)
XVIII	Heparin cofactor II	Heparin-dependent serine protease that inhibits thrombin	Deficiencies result in thromboembolism (see factor XVII)
XIX	Protein C	Inactivates Va and VIIIa	Mutations in *PROTC* lead to thrombophilia (see factor XVII)
XX	Protein S	Vitamin K-dependent cofactor of activated protein C	Mutations in *PROTS* result in thrombophilia (see factor XVII)

[a] Factor III is also used as a synonym for thromboplastin, a mixture of tissue factor (TF) and phospholipid.
[b] High-M_r kininogen.

Table 13.2 Clotting factors. ⏎

Classification of clotting factors

It is convenient to classify clotting factors into four discrete groups based on their action(s) in the clotting process:

1. zymogens
2. cofactor proteins
3. protease inhibitors
4. miscellaneous group

Zymogens (see *Margin Note 11.4*) only become active proteases when subjected to specific proteolytic cleavage(s), while **cofactor proteins** bind to zymogens and their protease products and increase the specificity and speed of the activation. Protease inhibitors inactivate the proteases after their roles in the clotting process are complete. The miscellaneous fourth group includes fibrinogen, factor XIII/XIIIa and von Willebrand factor (vWF). The first is hydrolytically cleaved to form the initial fibrin clot, which factor XIIIa, a transglutaminase, stabilises by catalysing the formation of covalent cross links between fibrin molecules. von Willebrand factor carries factor VIII (see below and *Section 13.9*) in the plasma and also anchors platelets to the endothelium of blood vessels.

Clotting pathways

The period before blood flow stops, the **bleeding time**, is normally two to six minutes, and ceases when a blood clot seals the lesion. Clotting begins with one of two pathways that both feed into a common final one:

- intrinsic
- extrinsic
- common (third) final

The intrinsic and extrinsic pathways each have several unique reactions but, in the end, both activate a final clotting stage, which is the formation of factor Xa (*Figure 13.16*).

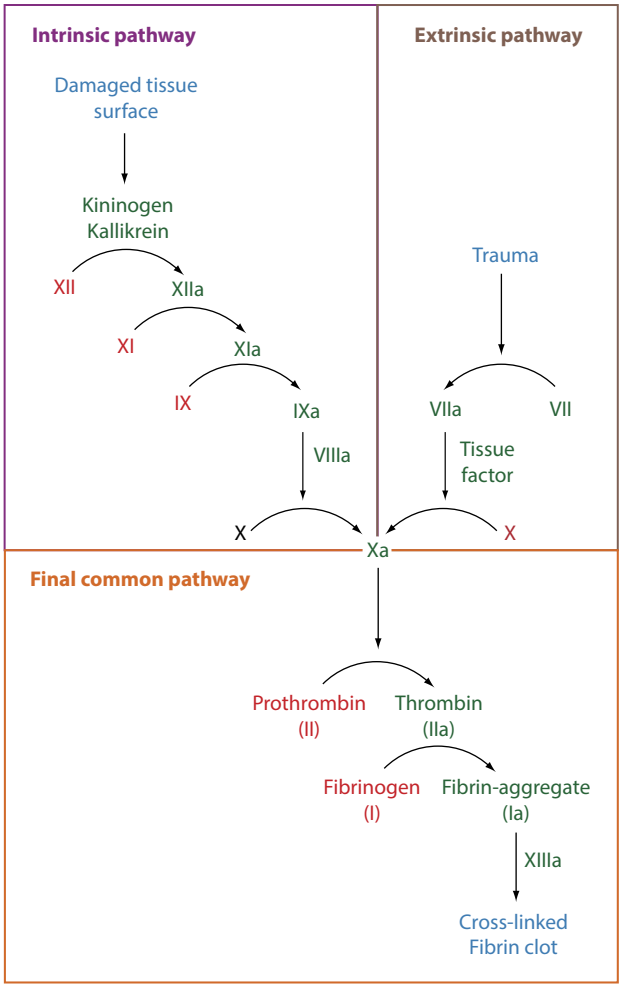


Figure 13.16 A simplified form of the blood-clotting cascade, consisting of the intrinsic, extrinsic and common final pathways. ⏎

Intrinsic pathway

The intrinsic pathway is a sequence of enzyme-catalysed reactions that becomes activated when a tissue is injured. This pathway can take several minutes for completion. Damage to the wall of a blood vessel or to its endothelium lining (*Chapter 14*) results in

the exposure of collagen to which platelets stick and then become activated. The platelets degranulate, releasing a range of highly bioactive substances, including serotonin (5-hydroxytryptamine or 5-HT; see *Figure 7.5*) and ADP, as well as several growth factors, for example platelet-derived growth factor (PDGF; *Figure 13.17*), which has a role in the subsequent healing of the wound. Serotonin is a powerful vasoconstrictor and causes blood vessels to constrict, temporarily limiting the blood supply to the damaged area until clotting has taken place. This is a temporary effect, and a clot must form soon. Platelet aggregation occurs as the released ADP activates other platelets and these, in turn, activate others to extend pseudopodia and become 'sticky' and clump with those already adhering to the damaged area. This forms a plug that prevents further blood loss. Fibrin is then deposited in the loose platelet plug, which traps more platelets as well as blood cells, strengthening the clot. This **clotting time** is usually 6 to 12 minutes.

The complex cascade of clotting reactions required to produce fibrin is initiated by the binding of factor XII (Hageman factor) to tissue collagen and anionic surfaces, which change its conformation to that of factor XIIa. This catalyses the conversion of prekallikrein to kallikrein. Kallikreins are serine endopeptidases that are widely distributed in mammalian tissues and body fluids. Plasma kallikrein, also known as kininogenin or Fletcher factor, cleaves bonds in kininogen to produce a varied group of polypeptides, which includes angiotensin, bradykinin, substance P and secretin. Kallikrein also activates the collagen-bound factor XII producing factor XIIa in a reaction that also requires high-M_r kininogen as a cofactor. Factor XIIa and high-M_r kininogen activate factor XI (plasma thromboplastin antecedent) to its active form, factor XIa. Factor XIa also binds to the exposed surfaces of the tissue where the injury has occurred. Here it catalyses the conversion of factor IX (Christmas factor) to factor IXa in a reaction that requires Ca^{2+}. The final reaction, unique to the intrinsic pathway, involves the conversion of factor X (Stuart factor) to Xa (*Figure 13.18*), which is its final product in a reaction catalysed by factor IXa. This reaction also requires Ca^{2+} and factor VIII, the antihaemolytic factor (*Figure 6.3*), which is carried in the circulation by von Willebrand factor/protein and found on the surface of aggregated platelets.

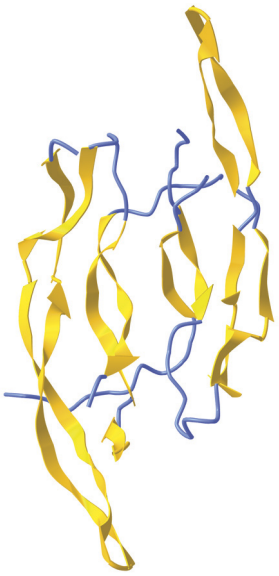


Figure 13.17 Molecular model of platelet-derived growth factor (PDGF). PDB file 1PDG.

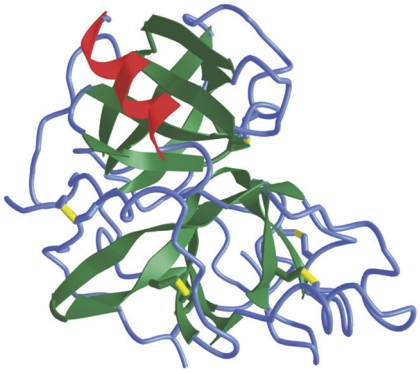


Figure 13.18 Molecular model of factor Xa. Two bound Na^+ are depicted as yellow spheres. PDB file 2Y5G.

Extrinsic pathway

The extrinsic pathway is so called because it requires the nonplasma protein, tissue factor (TF, formerly thromboplastin) to initiate the cascade. Tissue factor is an integral membrane protein found in many tissues but especially the walls of blood vessels, brain, lungs and placenta. The initial reaction is the conversion of factor VII (proconvertin) to factor VIIa in a reaction that requires Ca^{2+} and phospholipids released from the injured tissue or from the surface of aggregated platelets. Damaged tissues also release soluble TF, which, combined with factors VIIa (*Figure 13.19*) and IXa formed by the intrinsic pathway, directly activates factor X to Xa. The result is, like the intrinsic pathway, the production of factor Xa. Despite this common product, deficiencies in either pathway result in prolonged bleeding times.

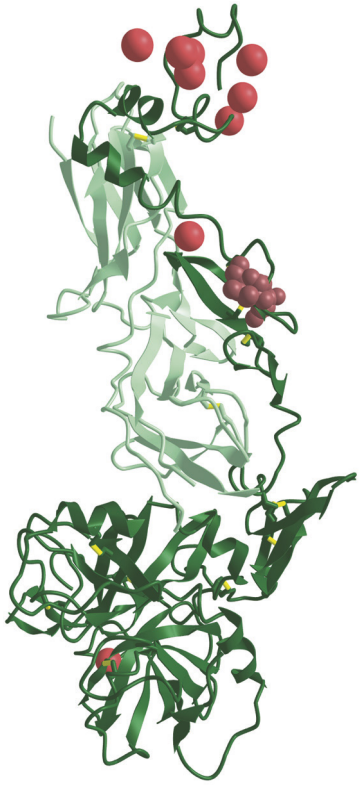


Figure 13.19 Molecular model of a soluble tissue factor (TF)–factor VIIa complex. The TF is illustrated in lighter green. The bound Ca^{2+} are shown in red. PDB file 2EC9.

Common (third) final pathway

Factor Xa formed by the intrinsic and extrinsic pathways catalyses the hydrolysis of prothrombin (factor II) to thrombin (IIa; *Figure 13.16*). There is an intermediate stage as prothrombin is first converted into prethrombin, which is then slowly converted to thrombin. Factor Xa by itself is a relatively slow activator of prothrombin, but its activity is enhanced about 20 000-fold in the presence of Ca^{2+}, factor V (proaccelerin) and negatively charged phospholipids, for example phosphatidylserine (*Figure 13.20*), which occur almost exclusively on the cytosolic sides of plasma membranes and so only contact

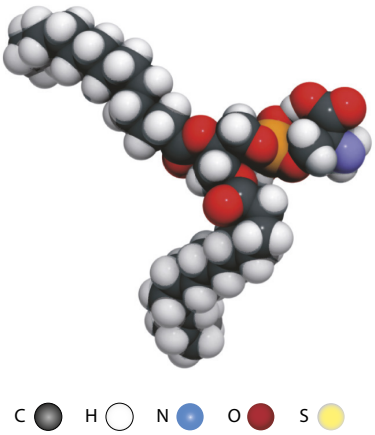


Figure 13.20 Computer drawn model of a phosphatidylserine.

the blood when the membrane is damaged. Hence clotting reactions are confined to sites of injury.

Thrombin is a serine protease that catalyses the hydrolysis of fibrinogen to fibrin (*Figure 13.21*), the final reaction of the clotting cascade. Fibrinogen forms 2–3% of the plasma proteins. It consists of three pairs of polypeptides and two pairs of oligosaccharides (*Figure 13.21*). Thrombin cleaves four peptide bonds in the fibrinogen molecule, releasing two A and two B fibrin peptides from the amino terminal ends of the fibrinogen polypeptides, changing the net charge from positive to negative. In the presence of factor XIIIa and Ca^{2+}, the fibrin monomers polymerise, producing a stable clot. The clots formed are strengthened by cross links formed by the transaminase factor XIIIa, which is also activated by thrombin. Within the clot, the platelets contract, reducing the clot to less than half its original size and one that is both tougher and more elastic, a process that draws the edges of the wound together and facilitates healing.

Margin Note 13.4 Fibrinogen

Fibrinogen is an **acute phase protein** (*Chapter 4*) whose concentration is substantially increased in certain clinical situations, such as acute inflammation caused by surgery, infections and myocardial infarction.

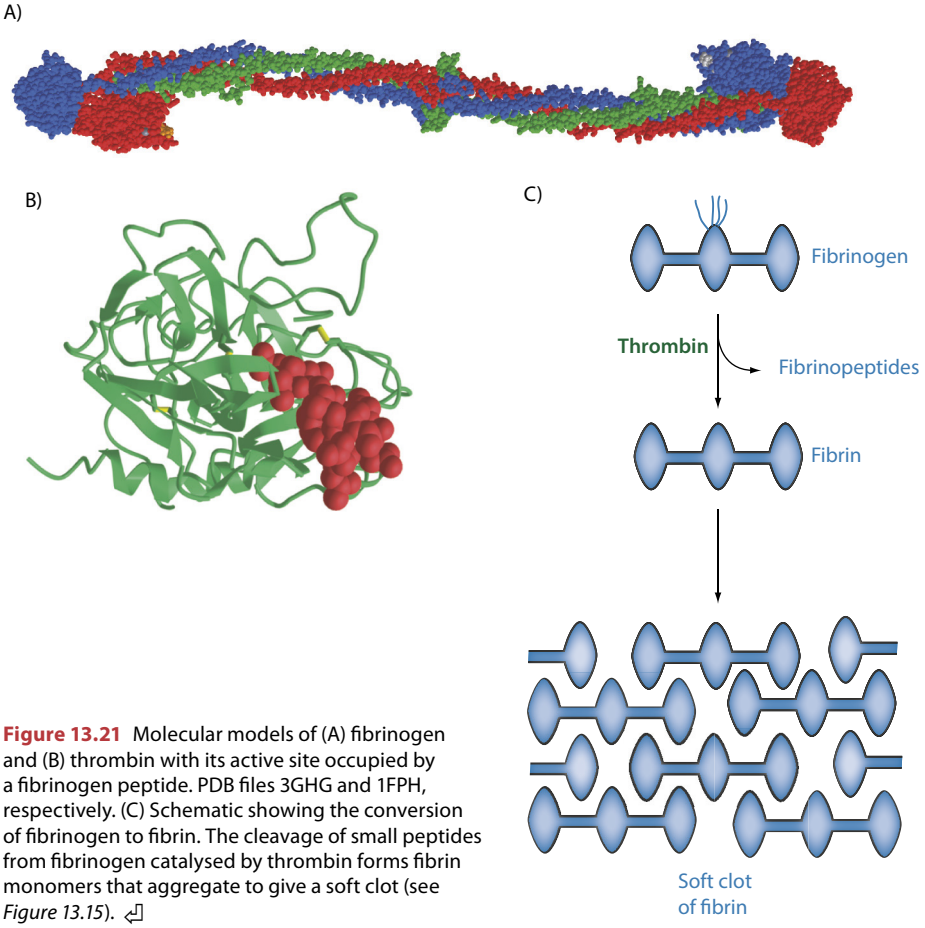


Figure 13.21 Molecular models of (A) fibrinogen and (B) thrombin with its active site occupied by a fibrinogen peptide. PDB files 3GHG and 1FPH, respectively. (C) Schematic showing the conversion of fibrinogen to fibrin. The cleavage of small peptides from fibrinogen catalysed by thrombin forms fibrin monomers that aggregate to give a soft clot (see *Figure 13.15*).

BOX 13.3 THROMBIN, PHYLLOQUINONE (VITAMIN K) AND ANTICOAGULANTS

Like many other blood-clotting factors, prothrombin is produced in the liver in a process that requires phylloquinone/vitamin K (see *Figure 10.23*). A lack of the vitamin results in the production of an abnormal prothrombin, which is activated by factor X at only about 1% of the normal rate. Native prothrombin contains 10 γ-carboxyglutamate residues. Phylloquinone/vitamin K is a cofactor for the enzymic conversion of the relevant glutamate to γ-carboxyglutamate residues in prothrombin (*Figure 13.22*). Dicoumarol and warfarin are competitive inhibitors of this process. Dicoumarol was first discovered in spoiled sweet clover (*Melilotus officinalis*) because it causes fatal haemorrhages in cattle, and warfarin was originally developed as rat poison. Agents such as warfarin are used for anticoagulant therapy. They are slow to produce effects, but they have long plasma half-lives and so can be active for days (*Chapter 14*). Warfarin is being gradually replaced by newer anticoagulants, such as apixaban, dabigatran and rivaroxaban (*Figure 13.23*). These anticoagulants do not require as much monitoring of the patient or dose adjustment but are as effective as warfarin as anticlotting agents; hence, their key advantage is convenience rather than improved safety.

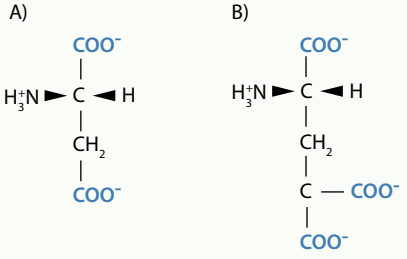


Figure 13.22 (A) Glutamate and (B) γ-carboxyglutamate.

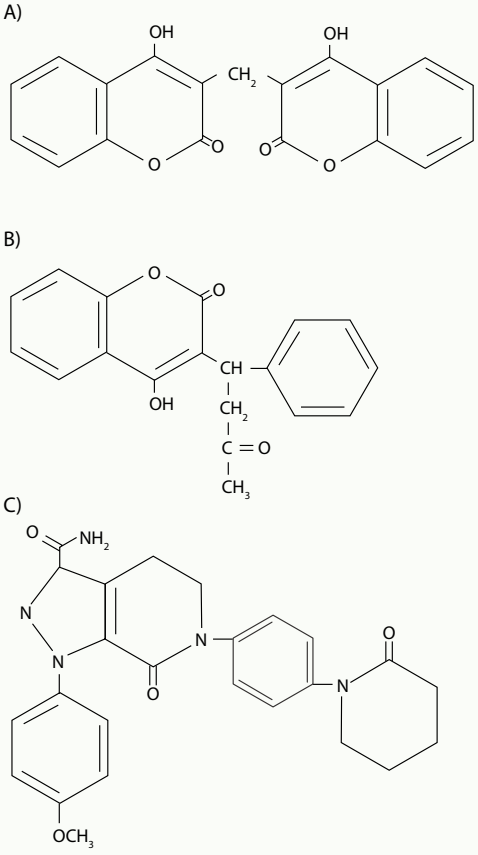


Figure 13.23 (A) Dicoumarol, (B) warfarin and (C) apixaban.

Prevention of clotting

It is clearly essential that clots do not form in the absence of injury. Several systems prevent this occurring. Although protein factors constantly circulate in the blood, they, of course, require specific activation before clotting can take place; also, the liver removes activated factors. This is, however, a relatively slow preventative measure. A more rapid mechanism is effected by α and β antithrombin IIIs, which are natural inhibitors of all the serine proteases of the clotting cascade but especially thrombin. Heparin, a sulphated polysaccharide found in the circulation, activates antithrombin III. Since β antithrombin III has the higher affinity for heparin, it is the major inhibitor *in vivo* (*Figure 13.24*). Patients with antithrombin III deficiency have an increased risk of thrombosis and resistance to the action of heparin. In addition, the plasma protein α_2-macroglobulin also inhibits the clotting cascade.

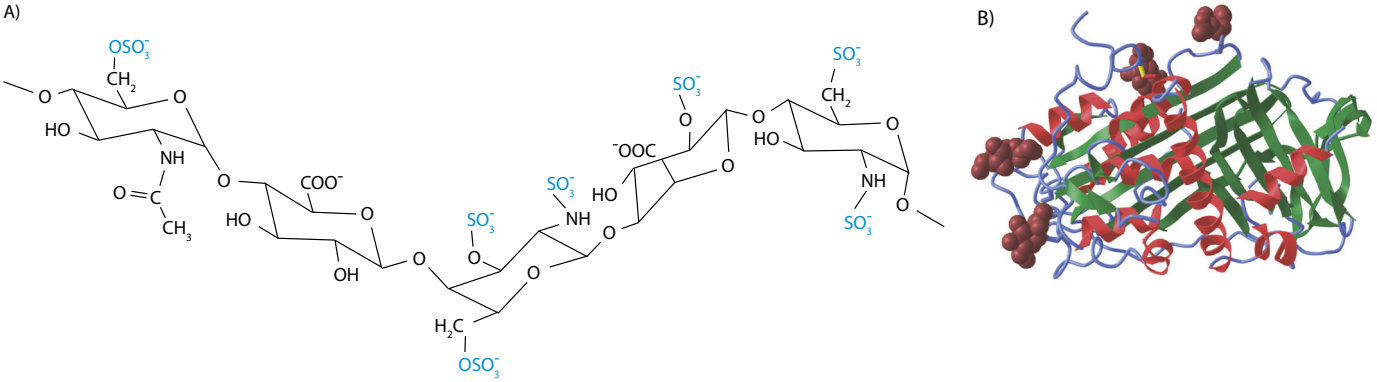


Figure 13.24 (A) A pentasaccharide portion of the anticoagulant polysaccharide heparin. The M_r of a complete molecule is approximately 17 000. (B) Molecular model of β antithrombin III. Its glycans are coloured brown. PDB file 1E05.

Dissolution of clots

Clots on the skin surface eventually scab and are largely removed by abrasion. However, internal clots are eventually destroyed by a process called fibrinolysis. Again, a cascade of reactions is involved. Fibrin activates plasminogen activator, a protein that, as its name implies, converts plasminogen to plasmin (*Figure 13.25*). Plasmin is a hydrolytic enzyme that catalyses the digestion of fibrin and dissolves the clot. Plasminogen activator has a high affinity for fibrin clots and it and a tissue-type plasminogen activator (t-PA) bind to the clot and activate plasminogen. In some instances (such as prolonged bleeding, shock and some types of cancer), plasminogen activator can be activated in the absence of fibrin. In these cases, plasminogen activators, such as streptokinase or t-PA, are given to patients to dissolve blood clots to try to reduce the damage caused by myocardial infarction (*Chapter 14*).

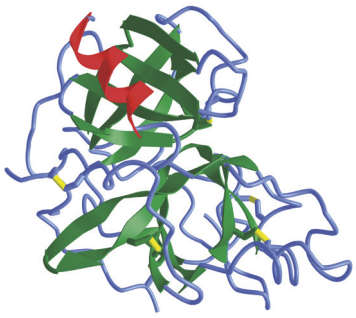


Figure 13.25 Molecular model of plasmin. PDB file 6Q1U.

13.5 METHAEMOGLOBINAEMIA

Methaemoglobinaemia (MetHb) is a group of disorders characterised by excessive amounts of methaemoglobin in the blood. It can be acquired or inherited. Acquired forms are most common. They occur in susceptible individuals on exposure to chemicals that oxidase Fe^{2+}. These include nitrites used in food preservation; some medications such as the topical anaesthetics benzocaine, lidocaine and prilocaine; and dapsone, which is used to treat some forms of dermatitis and fungal lung infections in human immunodeficiency virus (HIV)/acquired immunodeficiency syndrome (AIDS) patients, and by some antimalarial drugs (*Chapters 2* and *3*). Recreational drug use involving amyl nitrite (poppers) and nitrous oxide (laughing gas) can also be a cause.

Symptoms of acquired MetHb include fatigue, headaches, pallor and weakness. However, depending on its cause, acquired MetHb may depress the central nervous system (CNS), causing drowsiness, slow reflexes, slurred speech, nausea and vomiting, and metabolic acidosis (*Chapter 9*). In some cases, movements may become uncontrollable and there may be loss of consciousness.

Hereditary MetHb occurs in rare individuals with genetic deficiencies in NADH cytochrome b_5 reductase activity (see also *Margin Note 13.5*). All are inherited in an autosomal recessive manner. Four types, termed I–IV MetHbs, result from different mutations in the genes encoding the enzyme subunits. Type I is the most common subtype, with the enzyme deficiency occurring only in erythrocytes. In contrast, in type II MetHb, enzyme activity is deficient in all body tissues. Type III MetHb affects the haematopoietic

Margin Note 13.5 Haemoglobin M-Saskatoon and hereditary methaemoglobinaemia

Mutations in the genes for the α or β chains of Hb can also lead to a hereditary MetHb. If the mutations result in the substitution of tyrosine for histidine residues in or near the haem pocket, they produce HbM-Saskatoon and a haemoglobinopathy (*Section 13.7*). Haemoglobin M-Saskatoon has an increased propensity to form methaemoglobin rather than oxyhaemoglobin in the presence of dioxygen. The disorder is inherited in an autosomal dominant manner.

BOX 13.4 CYANOSIS: BLUE IN THE FACE/BLUE BABIES

Cyanosis (from the Greek *cyan*, blue) is the bluish complexion resulting from lack of oxygen in the circulating blood. It is most frequently observed under the nails, in lips and in the skin. Cyanosis occurs following an inadequate oxygen intake in the lungs or, for example, the stagnation of blood in the circulation during heart failure (*Chapter 14*), or as a result of a cyanotic congenital defect, an inborn cardiac condition that allows venous (deoxygenated) blood to mix with arterial (oxygenated) blood, producing a cyanosis. The most common heart defect leading to this condition is the tetralogy of Fallot or Fallot syndrome, named after the French physician of that name (Etienne-Louis Fallot, 1850–1911). It consists of four congenital heart defects (*Figure 13.26*):

1. an abnormal opening in the heart septum
2. pulmonary stenosis (a constriction of the pulmonary valve)
3. malposition of the aorta associated with the defect in the right ventricle (dextroposition)
4. hypertrophy of the right ventricle

The tissues of affected babies, 'blue babies', are deficient in oxygenated blood, cardiac murmurs are audible with a stethoscope and babies suffer dyspnoea, hypoxia, tachycardia, tachypnoea and seizures (*Chapter 14*). Growth may be retarded, and clubbing of phalanges may occur. Older children often assume a characteristic squatting position after exercise to alleviate the breathlessness induced by the hypoxia. Treatment is primarily supportive and palliative surgical procedures until the baby has grown sufficiently, usually aged about one year, to allow full corrective surgery.

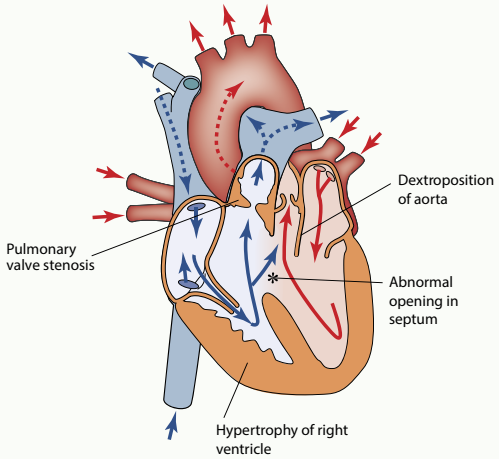


Figure 13.26 Illustration of the four congenital heart defects constituting the tetralogy of Fallot or Fallot syndrome.

system, that is, erythrocytes, leukocytes and platelets, while type IV resembles type I in affecting only erythrocytes. Homozygous patients suffering from types I, III and IV are cyanotic when concentrations of methaemoglobin exceed 1.5%. In type IV, the cyanosis is chronic. Type II MetHb particularly affects the CNS, leading to severe mental deficiency by nine months of age, which may be a consequence of the function of NADH cytochrome b_5 reductase in the desaturation of fatty acids.

An initial diagnosis of MetHb is made on the presence of cyanosis at birth and a low blood oxygen that does not improve with oxygen therapy combined with an absence of cardiac or pulmonary disorders. The diagnosis can be confirmed by measuring the methaemoglobin concentration and by assaying NADH cytochrome b_5 reductase activity. A diagnosis can be confirmed by a blood gas analysis (*Chapter 8*). Enzymatic and molecular genetic testing of both parents can assist in distinguishing type II from the other forms of MetHb, since it is difficult to discriminate between them until differences in cognitive development are apparent.

Treatment of both the acquired and the inherited form is generally with oxygen therapy and intravenous methylthioninium chloride, formerly called methylene blue, which reduces $Fe(III^{3+})$. Common side effects of methylthioninium chloride include headaches, vomiting, confusion, shortness of breath and high blood pressure and it is generally used when methaemoglobin concentrations exceed 30% or the cyanosis does not respond to oxygen therapy. Other treatments may include vitamin C, blood exchange transfusions and hyperbaric oxygen therapy. Outcomes for types I, III and IV are generally good with treatment. The severe MetHb of type II can usually be alleviated by doses of 200–500 mg/day of ascorbic acid but there is no effective treatment for the associated neurological disorder, and patients rarely live beyond infancy.

13.6 ANAEMIAS: LOW HAEMOGLOBIN CONCENTRATIONS

Anaemia (from the Greek *a* and *haima*, without blood) is the clinical condition whereby the amount of Hb in the blood falls below reference values for an individual in a given population (*Chapter 1*). It may be caused by insufficiencies of iron and some vitamins (*Chapter 10*) necessary for Hb synthesis, major blood loss or defective Hb molecules (**haemoglobinopathies**, see *Section 13.7*). The characteristic signs of anaemia are **tachycardia** (a fast heart rate), cardiac failure (*Chapter 14*), pallor and epithelial changes including atrophy of the tongue papillae, angular stomatitis, brittle hair and nails, and spoon-shaped nails. Other signs specific to the type of anaemia may also be present. However, anaemic patients may be asymptomatic, even when the anaemia is quite severe, or may present with various nonspecific symptoms, such as angina on physical effort, breathlessness, fatigue, headaches or palpitations (*Chapter 14*). A rapid onset of anaemia tends to cause more symptoms than a slow onset, and the elderly tolerate anaemia less well than the young when the normal compensation of increased cardiac output is impaired.

The body responds to anaemia with a variety of physiological responses. For example, the heart responds with an increased stroke volume and tachycardia to increase its output (*Chapter 14*). The oxygen binding curve of Hb may be modified by the production of 2,3 bisphosphoglycerate (BPG), which increases the release of O_2 from oxyhaemoglobin to the tissues. In iron deficiency anaemia, the concentration of BPG can increase by 40–75%. The reduced amount of Hb in anaemia is usually accompanied by a decreased erythrocyte count (RBC, echoing their alternative name of red blood cells) and packed cell volume (PCV), which are usually determined in hospital laboratories using an automated blood analysis system (*Figure 13.6*). These types of systems can determine all haematological indices, for example the Hb concentration, PCV and RBC, for numerous samples rapidly and efficiently. The mean cell volume (MCV), mean cell Hb (MCH) and mean cell Hb concentration (MCHC) can be derived from these measured values (*Box 13.5*). Any investigation of anaemia should also consider the reticulocyte, leukocyte and platelet counts and any abnormal morphology observed in a blood and/or marrow specimen when examined by microscopy.

Types of anaemias

The three major types of anaemias are:

1. normocytic
2. microcytic
3. macrocytic

BOX 13.5 HAEMATOLOGICAL INDICES

The measured erythrocyte values for haemoglobin (Hb) concentration, packed cell volume (PCV) and erythrocyte count (RBC; see *Table 13.1*) allows four haematological indices to be determined:

1. PCV
2. mean cell volume (MCV)
3. mean cell Hb (MCH)
4. mean cell [Hb] (MCHC)

The PCV can be measured using an automated cell counter (*Box 13.2*). In teaching exercises, it can be determined using a haematocrit as shown in *Figure 13.27*. It can also be derived as the product of the MCV and RBC. The MCV is obtained by dividing the PCV by the RBC. The MCH can be calculated by dividing the Hb concentration of whole blood by the RBC. The MCHC can be obtained by dividing the Hb concentration by the PCV. The normal values for these indices are given in *Table 13.3*.

	Males	Females
[Haemoglobin] (g/L)	135–175	115–155
PCV (haematocrit) (%)	40–52	36–48
MCV (fL)	80–95	80–95
MCH (pg)	27–34	27–34
MCHC (g/L)	200–350	200–350
[Serum iron] (μmol/L)	10–30	10–30
Total iron-binding capacity (μmol/L)	40–75	40–75

MCH, mean cell haemoglobin; MCHC, mean cell haemoglobin concentration; MCV, mean cell volume; PCV, packed cell volume (*Box 13.5*).

Table 13.3 Normal blood index values.

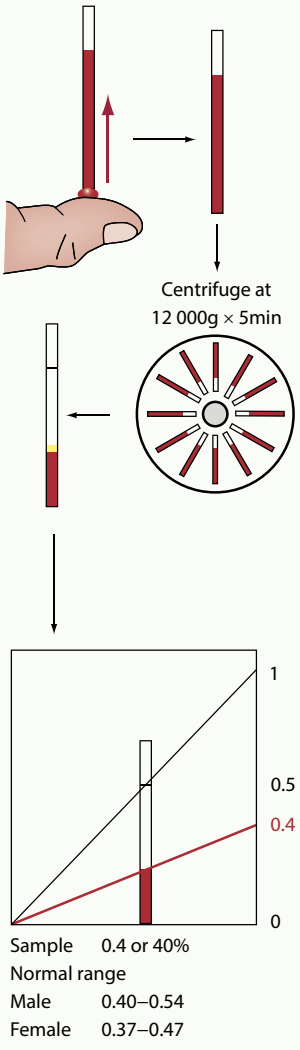


Figure 13.27 Using a haematocrit to determine the packed cell volume (PCV) after collecting blood from a thumb prick into a capillary tube.

These are classified largely in terms of erythrocyte indices, especially their MCV values. Normocytic anaemias, that is, with normal-sized erythrocytes, are associated with acute blood loss and a variety of disease states. However, the microcytic and macrocytic anaemias are associated with observable changes to the sizes of erythrocytes in the blood sample and distinctive changes to the bone marrow appearance (*Table 13.4*). Thus, the classification of anaemias starts from routine haematological investigations. Their diagnosis involves taking a medical history and a clinical investigation of the patient, especially, of course, examining the blood and marrow to determine any changes in erythrocyte colour, size and shape.

Erythrocyte appearance	Diameter (μm)	Mean cell volume (fL)	Causes
Normocytic and *normochromic* (normal size and colour)	7	76–96	Acute blood loss, infection, collagen disease, malignancy, endocrine disease, chronic disease
Microcytic and *hypochromic* (small cells, pale due to reduced haemoglobin content)	<7	<80	Iron deficiency, thalassaemias, sideroblastic anaemia, chronic disease
Macrocytic (large cells; oval or round in shape)	>9	96	Deficiencies of vitamins B_{12} or folate (oval), alcoholism (round) or liver disease (round)

Table 13.4 A classification of anaemias.

Normocytic anaemias

Normocytic anaemia is characterised by the presence of erythrocytes that are normal in colour (normochromic), size and shape, but are insufficient in number to satisfy metabolic requirements. It is among the most common forms of anaemia and its incidence increases with age. It is often only discovered by routine laboratory tests. Normocytic anaemia can be congenital, iatrogenic or the result of trauma, such as bleeding from an injury and the chronic bleeding associated with a gastrointestinal problem. It most often results from some or another underlying clinical condition, when it is known as anaemia of chronic disease (ACD) or anaemia of inflammation. These conditions include kidney disease; disorders of the immune system, such as rheumatoid arthritis, lupus erythematosus and inflammatory bone marrow disorders (*Chapter 5*); infections (*Chapter 2*); cancers and their treatments (*Chapter 17*); and malnutrition (*Chapter 10*). Heavy menstrual bleeding and, indeed, pregnancy can also give rise to normocytic anaemia.

Symptoms of normocytic anaemia are slow to develop. Its initial signs are usually dizziness, fatigue, pallor and a shortness of breath. Given the vagueness of these symptoms and because normocytic anaemia is frequently associated with a chronic condition, it can be difficult to distinguish the anaemia symptoms from those of the underlying problem. Normocytic anaemia is diagnosed by taking a thorough clinical history and can be identified with a blood examination, such as a complete blood count (CBC), especially to determine the number, colour, size and shape of the erythrocytes. The anaemia is marked by seemingly healthy, normal-looking erythrocytes but which are reduced in number (*Table 13.4*). In some cases, a bone marrow biopsy may also be performed. When a genetic link is suspected, investigations of family members are necessary. If the anaemia is linked to a chronic clinical condition, then managing this must be the first priority with, for example, anti-inflammatory medications for rheumatoid arthritis or antibiotics to combat infections. If the anaemia is due to insufficient dietary iron, then supplements of ferrous sulphate are appropriate. Severe normocytic anaemia is treated with blood transfusions, but care must be exercised to prevent iron overload. The use of recombinant human erythropoietin (EPO; *Figure 13.28*), a cytokine that targets erythrocyte precursors in response to O_2 tissue requirements, to increase bone marrow erythrocyte production may also be considered. Acute normocytic anaemia should generally show increases in erythrocyte numbers within a few weeks of beginning treatment. The improvements are generally long-lasting and do not require prolonged therapy. In contrast, chronic normocytic anaemia may require prolonged treatment of the underlying cause to restore the erythrocyte count and prevent any recurrence.

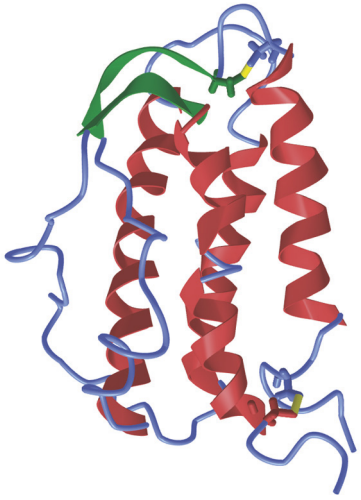


Figure 13.28 Molecular model of erythropoietin (EPO). PDB 1BUY.

Microcytic anaemias

The major causes of microcytic anaemias are iron deficiency, thalassaemias (*Section 13.7*), sideroblastic anaemia and ACD. Iron is difficult to absorb because of problems connected with the low solubility of its salts, its oxidation state and interaction with other components of the diet (*Chapter 11*). Loss of iron also occurs in haemorrhage and menstruation (*Box 13.1*).

Iron deficiency anaemia, the most common cause of anaemia worldwide, shows a number of characteristic features. The erythrocytes are microcytic (*Figure 13.29*), with an MCV of less than 80 fL, and hypochromic, the MCH being less than 27 pg. There are variations in cell sizes, **anisocytosis** and **poikilocytosis** that is, abnormal shapes, and a reduced reticulocyte count. There are also changes in the bone marrow, for example erythroid hypoplasia and decreased iron deposits. The serum iron concentration decreases and the serum iron-binding capacity increases compared with their normal values (*Table 13.3*). The transferrin saturation, that is, the proportion of serum iron to its total iron-binding capacity, is less than 19% compared with a more usual 30% or so. The concentration of serum ferritin resulting from cellular degradation is regarded as the most reliable measurement of anaemia. Patients may also show an abnormal thyroid hormone metabolism (*Chapter 7*), depressed muscle function and an impaired ability to maintain body temperature.

The underlying cause of the iron deficiency should be identified by taking a careful clinical history of the patient and applying appropriate clinical tests. To counteract any deficiency, 600 mg of ferrous sulphate is given orally each day. If there are side effects, such as nausea, diarrhoea or constipation, then ferrous gluconate may be substituted. Failure to respond to the treatment may be due to lack of patient compliance, continuing haemorrhage, severe malabsorption or another cause of the anaemia. It may be necessary to give iron parentally if absorption is defective, for example in patients with ulcerative colitis or Crohn disease (*Chapter 11*).

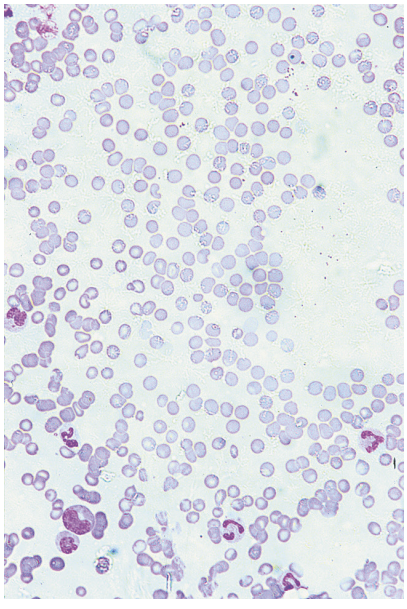


Figure 13.29 Photomicrograph of a peripheral blood smear from a patient suffering from severe iron deficiency. The cells are microcytic and hypochromic.

Macrocytic anaemia

Macrocytic anaemias are characterised by the presence of anaemia and erythrocytes of variable shapes but with diameters exceeding 9 μm and MCVs characteristically greater than 96 fL (*Table 13.4*). The condition may be caused by certain liver diseases, including alcoholism, that produce large rounded erythrocytes or by megaloblastic anaemia, which is associated with enlarged oval erythrocytes and by the presence of the erythrocyte precursors erythroblasts (megaloblasts) in blood. The increased proportion of immature forms of all cell lines reflects the premature death of cells in the process of development (*Figure 13.30*). The cells are large, although there is a substantial variation in size, and they have large, immature nuclei. The problems stem from the inability to synthesise deoxythymidine monophosphate from methylated deoxyuridine monophosphate using a methyl group supplied by the vitamin B_9 (folate)-based coenzyme methylene tetrahydrofolate polyglutamate in a reaction catalysed by thymidylate synthase (see *Figure 17.33*). A deficiency of folate (*Chapters 10* and *11*) impairs the reaction. A deficiency of vitamin B_{12} (cobalamin) also reduces its supply by slowing the demethylation of methyl tetrahydrofolate. Thus, deficiencies of folate or cobalamin (*Chapters 10* and *11*) or other defects, for example genetic ones that affect DNA biosynthesis in the bone marrow, produce an asynchrony between nuclear and cytoplasmic development. This delays the maturation of blood cells and results in megaloblastic anaemia.

The causes of folate deficiency are nutritional (for example a poor intake of green vegetables such as broccoli and spinach and of offal), alcohol excess, cancer, its excessive utilisation in pregnancy and lactation, and the use of antifolate drugs such as methotrexate, phenytoin and pyrimethamine. The clinical manifestations of folate deficiency are megaloblastic anaemia with a serum folate concentration lower than its reference value of 4–18 μg/L.

Pernicious anaemia is a macrocytic anaemia disease that predominantly affects the elderly with 1 in 8000 of the over-60-year-old population affected in the UK. Its cause is a reduced intake/absorption of vitamin B_{12} (cobalamin), although it also seems to be associated with certain autoimmune diseases, such as thyroid and Addison diseases (*Chapter 7*). The rather complicated absorption of vitamin B_{12} is described in *Chapter 11*. It is then transported in the blood bound to transcobalamin (*Figure 13.31*). An atrophy of the gastric mucosa and consequent reduced production of HCl and gastric intrinsic factor (GIF) leads to its malabsorption and pernicious anaemia. Cytotoxic immunoglobulin G (IgG) antibodies directed against gastric parietal cells and/or against GIF are found in the sera in about 90% of individuals with pernicious anaemia. In a majority of these individuals, the antibodies are also present in the gastric juice and prevent or inhibit the absorption of the vitamin B_{12}–GIF complex.

The onset of the disease is insidious with progressively increasing symptoms of anaemia. Patients show **achlorhydria** (low or absence of HCl secretion) and a lack of secreted GIF. There may be jaundice because of excessive breakdown of Hb and because erythropoiesis in the bone marrow is deficient. The serum bilirubin may be increased and the serum vitamin B_{12} concentration is usually considerably below its physiological value of approximately 160 ng/L. However, the polyneuropathological symptoms make it essential that treatment is not delayed, as they can become irreversible. Patients present with symmetrical paraesthesia in the fingers and toes, an early loss of vibration sense and ataxia. Paraplegia may result.

The bone marrow can be investigated by microscopy. If a larger sample of bone marrow is required, the posterior iliac crest is used but a longer and wider needle is used to obtain a 'core' of bone. This core is fixed and decalcified over several days and then stained for microscopy. It is necessary to distinguish pernicious anaemia from other causes of megaloblastic anaemia, such as folate deficiency, because this will affect treatment. However, this is usually clear from the blood concentrations of these two vitamins.

Folate deficiency can be corrected by giving 5 mg of folic acid daily, and this usually produces a rapid haematological response. Prophylactic folate is recommended for all women during pregnancy and especially for women who have had a previous child with a neural tube defect. Vitamin B_{12} deficiency is treated by intramuscular injections of the pure vitamin, to a total of 6 mg over a period of three weeks. Oral administration is obviously ineffective in pernicious anaemia because of the lack of GIF. A maintenance dose of 1 mg every three months is then given for the rest of the patient's life. Clinical improvement may occur within a few days (provided that the neuropathy has not been long-standing) and a reticulocytosis is observed a few days later.

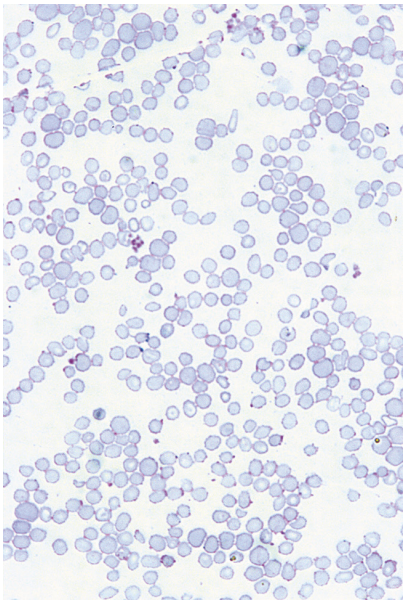


Figure 13.30 Photomicrograph of a marrow smear from a patient with megaloblastic anaemia. The mean cell volume is over 95 fL and the macrocytes are typically oval. The bone marrow is usually hypercellular and the normoblasts (erythroblasts) are large and show failure of nuclear maturation with an open, fine, primitive (stippled) chromatin pattern. Giant and abnormally shaped cells called myelocytes are present. ⏎

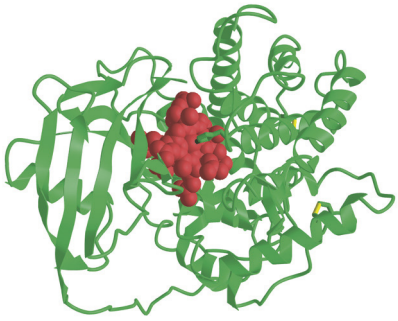


Figure 13.31 Molecular model of transcobalamin–vitamin B_{12}/cobalamin, with the vitamin highlighted in red. PDB file 2BB5. ⏎

13.7 HAEMOGLOBINOPATHIES

Haemoglobinopathies are clinical conditions resulting from mutations (*Chapter 15*) in the globin genes of Hb. Over 1300 mutations in the genes associated with Hb have been reported. If any of the bases in the DNA are changed even by one, then a modified Hb molecule may be produced (or no Hb at all). The consequences vary: they can be negligible, severe or fatal. Mutations are inherited and, if the disease is not fatal, then the disease symptoms will be inherited too. The severity of the disease may depend on whether one or both copies of the gene in question carry the mutation, that is, is the individual homozygous or heterozygous (*Chapter 15*). The mutations involved in haemoglobinopathies include point mutations (the largest group), which substitute one amino acid residue for another; insertions or deletion of one or more residues; drastic changes caused by frameshift mutations (*Margin Note 13.6*); and alterations in the lengths of the polypeptides that produce or destroy stop codons.

Margin Note 13.6 ⏎ Frameshift mutations

In the genetic code, sequences of three nucleotide bases or **codons** code for each amino acid residue. A change in one base may cause the incorporation of a different (or 'wrong') residue, such as occurs in HbS. However, if one or more bases is/are lost or added, then the reading frame of the sequence is shifted. Instead of a single amino acid residue being changed, a totally new sequence may be produced, which may result in the production of a different polypeptide/protein or often none at all, depending upon where in the base sequence the frameshift occurred.

BOX 13.6 NOMENCLATURE OF MUTANT HAEMOGLOBINS

Two nomenclature systems are in use to describe haemoglobins (Hbs), which may be a little confusing. Originally, normal adult Hb was called haemoglobin A (HbA [A for adult], $\alpha_2\beta_2$). Fetal Hb was haemoglobin F (HbF, $\alpha_2\gamma_2$), sickle cell Hb was HbS and so on. However, as more Hb mutations were identified, their number exceeded that of the letters in the alphabet. Subsequently, Hbs began to be named after the geographical location where they were first discovered, for example Hb Dakar, Hb Lepore and Hb Sydney. An additional complication occurs when newly discovered Hbs have the same characteristics as a 'letter' Hb and both nomenclatures may be combined, as in HbM-Saskatoon (see *Margin Note 13.5*). It is appreciated that this nomenclature is not perfect but it is too difficult to change now.

The precise mutation can of course be described in terms of the base change(s) when as is usual the sequence of the gene is known, but this is a little cumbersome for everyday use. For example, Hb Sydney is caused by a GTG to GGG mutation at position 67 in the gene for β globin, causing an amino acid residue change of valine to alanine. This results in unstable Hb with poor haem-binding properties, resulting in mild haemolysis. *Table 13.5* gives some examples of mutant Hbs.

Haemoglobin	Codon change	Amino acid residue change	Comments
Torino	TTC to GTC	Phe to Val	α chain (43); decreased O_2 affinity, unstable
Ann Arbor	CTG to CGG	Leu to Arg	α chain (80); unstable
Bibba	CTG to CCG	Leu to Pro	α chain (136); dissociates
M-Iwate	CAC to TAC	His to Tyr	α chain (87); forms methaemoglobin, benign cyanosis in heterozygotes
Constant Spring	UAA to CAA	STOP to Gln	α chain mutation of the chain termination codon (142) gives extended α chain
S	GAG to GTG	Glu to Val	β chain (6); cells sickle, forms fibrils
C	GAG to AAG	Glu to Lys	β chain (6); enhances sickling when with sickle cell haemoglobin
St Louis	CTG to CAG	Leu to Gln	β chain (28); Fe readily oxidised, polar group in haem pocket, increased O_2 affinity, unstable
Seattle	GCC to GAC	Ala to Asp	β chain (70); decreased O_2 affinity, unstable
Gun Hill	–	Deletion	β chain (91–95); increased O_2 affinity, unstable
M-Saskatoon	CAT to TAT	His to Tyr	β chain (63); forms methaemoglobin, benign cyanosis in heterozygotes

Table 13.5 Examples of the many known mutations in haemoglobin genes. ⏎

In normal adult humans, there are two α and one β globin genes (*HBA1* and *HBA2*, both at locus 16p13.3, and *HBB* at 11p15.4), coding for polypeptides/subunits of 141 and 146 amino acid residues, respectively, which form HbA, $\alpha_2\beta_2$. In diploid cells there are, naturally, four α and two β genes. Each of these genes has two introns (*Margin Note 13.7*).

If there is a mutation, it may have been inherited from one or both parents, giving a heterozygous or homozygous condition, respectively. A mutation in an α gene tends to have less serious consequences than those in a β gene because there may still be nonmutated copies of the α gene present. Nevertheless, even small changes in the structure of Hbs can sometimes have disastrous clinical effects. Given that there are separate genes for the α and the β globins, they usually only affect one type of subunit (*Table 13.5*).

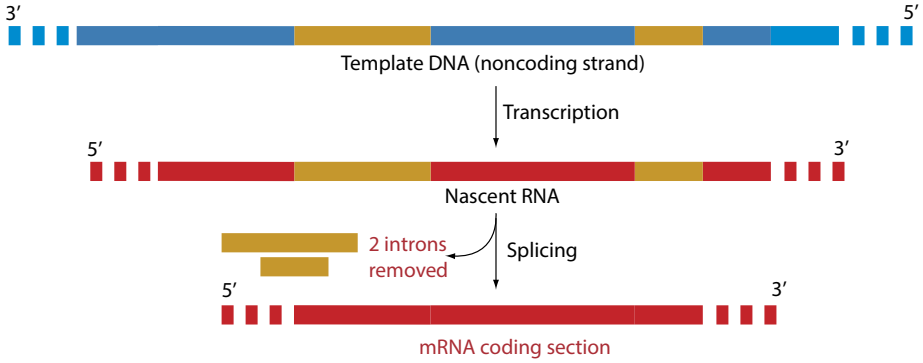


Figure 13.32 Outline of the transcription of a putative gene containing three exons and two introns and the splicing of the resulting nascent RNA molecule. ⏎

Margin Note 13.7 ⏎ Introns, exons and splicing

Most genes in eukaryotic cells are not continuous but are arranged in sections along the DNA in a chromosome. The coding or expressed sequences are called **exons**, and the noncoding portions between them are termed intervening sequences or **introns**. To synthesise messenger RNA that can be translated to produce a polypeptide, a nascent RNA transcript of the DNA is made, that is, one with both the exons and the introns transcribed. The introns are then excised and the newly formed ends joined together in a process called **splicing** (*Figure 13.32*). This is a normal part of the processes that produce mRNA molecules, which can be translated by ribosomes to produce native polypeptides in eukaryotic cells.

BOX 13.7 IDENTIFYING MUTATIONS IN HAEMOGLOBIN GENES

Originally, many of the haemoglobins resulting from mutations were identified by their different mobilities during electrophoresis (*Figure 13.33*) and peptide mapping (*Box 13.8*). However, modern DNA techniques, such as the polymerase chain reaction (PCR) and next-generation gene sequencing (*Chapters 4* and *17*), means they can be directly identified quickly and cheaply. The ability to make DNA probes specific for α or β chains mean it is possible to identify which mRNAs are being produced and identify any mutations present.

Different *clinical* variants can be understood at the *molecular* level. For example, in so-called 'haemoglobin H disease', it has been shown that there is only one of the four possible α globin genes present and functioning, so that only 25% of the normal amount of α globin mRNA is produced. The mutation causing this situation is a deletion not a point mutation.

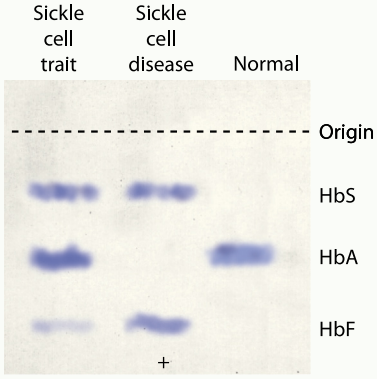


Figure 13.33 Use of electrophoresis to identify mutant haemoglobins with different charges from normal adult haemoglobin (HbA). For example, sickle cell haemoglobin (HbS) migrates more slowly towards the positive electrode because a glutamate residue (negatively charged) in its β chain has been replaced with a valine residue (zero charge). Hence, the whole molecule of HbS has two fewer negative charges than HbA. HbF, fetal haemoglobin. ⏎

Given that most mutations are harmless, they do not lead to a haemoglobinopathy. For example, mutations distant from the haem-binding cleft or from the regions of subunit contact may have little effect on the properties of the Hb. However, mutations may change the shape of the globin subunit(s) or the binding of the haem groups or even prevent globin synthesis, all with severe clinical consequences. To function properly, the four subunits in the Hb molecule must fit together tightly but still produce a molecule that is flexible. The regions of contact have been conserved in evolution and are essential for normal functions, such as the cooperative binding of O_2 (*Figure 13.11*). Thus, mutations can upset the delicate balance of interactions between the amino acid residue side chains, with several consequences. The molecule may dissociate upon deoxygenation and, in some cases, the monomers may precipitate in the erythrocytes, reducing O_2 affinity. Microscopically, the denatured and precipitated Hb can be seen as Heinz bodies (*Margin Note 13.8*). A deletion of one or more amino acid residues or substitution mutations can produce this effect, as in Hb Leiden and Hb Philly, respectively. There may also be plasma membrane damage, with intravascular haemolysis, anaemia, reticulocytosis

and splenomegaly (enlargement of the spleen) as consequences. In other cases, a small change in the regions that bind the haem groups may make the pockets slightly less hydrophobic so that it does not bind appropriately and, again, the denatured Hb can precipitate as Heinz bodies. Hence, only two of the four subunits may have haem groups. In other cases, the change in the pocket allows the iron to become oxidised to Fe^{3+} (methaemoglobin), which will not bind O_2 as described in *Section 13.3*. The resulting condition is referred to as methaemoglobinaemia and patients become cyanosed because they lack oxygen (*Section 13.5*). This chapter will concentrate on two haemoglobinopathies (see also *Margin Note 13.5*):

1. sickle cell anaemia
2. thalassaemias

Sickle cell anaemia

Sickle cell anaemia (SCA) or sickle cell disease (SCD) was first described in an African American patient in the USA in 1904, a time when little was known of the structures of proteins. The patient presented with severe pain, and a microscopic examination of a blood sample showed sickle-shaped erythrocytes (*Figure 13.35*). The mutated gene for sickle cell Hb (HbS) differs from that of β globin HbA by a single point mutation. Given that one β globin gene is inherited from each parent, the condition may therefore be homozygous (Hb^SHb^S) or heterozygous (Hb^AHb^S). The mutation occurs at codon 6 and substitutes a thymine for an adenine base leading to the acidic, hydrophilic glutamate residue being replaced by a hydrophobic valine in the β globin subunit *(Figure 13.36)*.

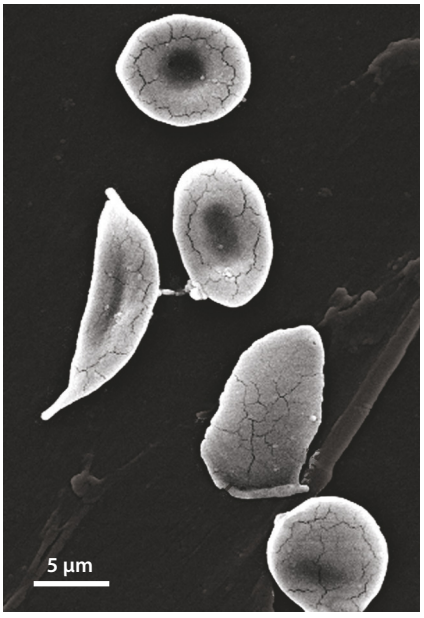


Figure 13.35 Electron micrograph showing sickle erythrocytes. Courtesy of the Centers for Disease Control and Prevention/Dr J.H. Carr.

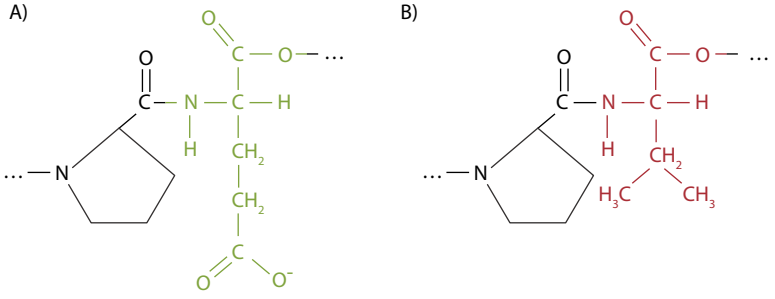


Figure 13.36 Dipeptides consisting of (A) –Pro–Glu– and (B) –Pro–Val– residues.

Margin Note 13.8
Heinz bodies

Heinz bodies or Heinz–Ehrlich bodies were first reported in 1890 by the German physician Robert Heinz (1865–1924) as inclusions in the erythrocytes of some patients with haemolytic anaemia. They are known to be aggregates of denatured, precipitated Hb, which associate with the erythrocyte membrane (*Figure 13.34*). This causes the erythrocyte to become misshaped and leads to anaemia.

The bodies are best seen when the blood is stained with crystal violet. Heinz bodies are associated with certain types of hereditary haemolytic anaemia, for example haemolytic anaemia of infancy. Oxidative damage to Hb by a number of toxic chemicals, including nitrobenzene, diphenylamine, naphthalene and hydroxylamine, and a number of its derivatives can also result in the formation of Heinz bodies. They also have iatrogenic causes and may result from sensitivity to some drugs, for example primaquine especially in glucose-6-phosphate dehydrogenase deficiency (*Section 13.8*) and sulphonamides and can appear after a splenectomy.

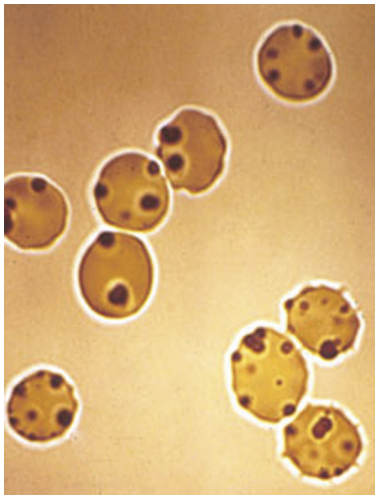


Figure 13.34 Heinz bodies in erythrocytes are formed from denatured, precipitated haemoglobin. Courtesy of Dr Ian Quirt, Department of Medical Oncology and Hematology, Princess Margaret Hospital, Toronto, Canada.

The Hb molecule is now somewhat more hydrophobic or 'sticky' in two places on its surface (*Margin Note 13.9*) because of the two β globins present. The sticky patches become more exposed in the deoxygenated state when the conformation of HbS changes as it releases its O_2 in the tissues. The deoxygenation of the erythrocyte results in β^S globins of different HbS deoxygenated molecules hydrophobically interacting, which initiates the formation of an HbS polymer. The speed of polymerisation is proportional to the concentration of HbS, but polymers can build up rapidly to form extended fibres, causing the erythrocyte to adopt a rigid, sickle morphology. The flexibility of the deformed erythrocytes is therefore much reduced compared with normal ones and they are less able squeeze through the capillaries and so block them. This leads to haemostasis, anoxia and severe pain and, because the sickling occurs as a result of changes occurring in deoxyhaemoglobin, the effects are exacerbated and more cells become sickle shaped. Damage to the vascular endothelium (*Chapter 14*) can increase superoxide formation and oxidative stress and inflammation. The life of a typical erythrocyte is reduced from 120 to about 10–12 days in SCA patients: the abnormal cells are destroyed in the spleen and consequently anaemia ensues.

The mutation that causes HbS production is not the only one that leads to sickling of erythrocytes, but the many other variants are rather rare. The second most common of these in Black populations occurs in haemoglobin C (HbC), in which a lysine residue replaces glutamate at position 6 in the β chain. Haemoglobin C is rather insoluble and crystals of it can sometimes be seen in peripheral blood smears. Heterozygotes for HbC are asymptomatic but homozygotes have a mild haemolytic anaemia.

The coincident distribution of the genes for HbS and HbC means that heterozygotes for both mutant Hbs are not uncommon, producing HbSC disease. This is milder than 'true' SCA but patients can show practically all the same complications. Furthermore, it is symptomatic in the heterozygous state. There can also be co-inheritance of the SCA and thalassaemia genes, which generates a wide spectrum of clinical symptoms whose severity depends on the type of thalassaemia mutation (see below).

At low O_2 concentrations, HbS has a normal oxygen binding curve (*Figure 13.11*), but at the high concentrations of oxygen in erythrocytes of a homozygote patient with SCA the affinity is decreased. The resulting shift in the oxygen dissociation curve to the right means a greater proportion of O_2 is released and ameliorates the effects of the anaemia. Indeed, the amino acid substitution that causes the condition does not affect the structure of the oxygen binding site or the ability of the molecule to bind and carry oxygen.

Margin Note 13.9
Why the surface?

The codon preceding position 6, that is, 5, codes for a proline residue (see *Figure 13.36*). Prolyls allow polypeptides to make tight turns, which in soluble proteins contributes to their globular conformations. Thus, they are often found on the surfaces of the molecules. Hence the following mutation changes the surface properties of the Hb molecule, that is, HbA to HbS:

…Pro–Glu…→…Pro–Val…

BOX 13.8 SICKLE CELL ANAEMIA: A MOLECULAR DISEASE

The identification of the precise mutation in sickle cell anaemia (SCA) was a significant historical step in the understanding of molecular diseases. Vernon Ingram (1924–2006) applied peptide mapping, sometimes called fingerprinting, in 1954 to the study of haemoglobin (Hb) and sickle cell Hb (HbS). The Hbs were digested into small peptides using the protease trypsin, and these were then separated in two dimensions on large sheets of chromatography paper using electrophoresis in one dimension, followed by chromatography in the second. It was found that the maps of Hb and HbS differed in the position of just one peptide (*Figure 13.37*), which was identified as being due to the change from a glutamate to a valine residue in the β globin as described in the main text. The determination of the change in Hb in SCA patients as an adenine to thymine mutation at codon 6 of the β globin gene was a major step in the understanding of genetic diseases. The condition was the first one to be referred to, by Linus Pauling (1901–1994), as a *molecular disease*.

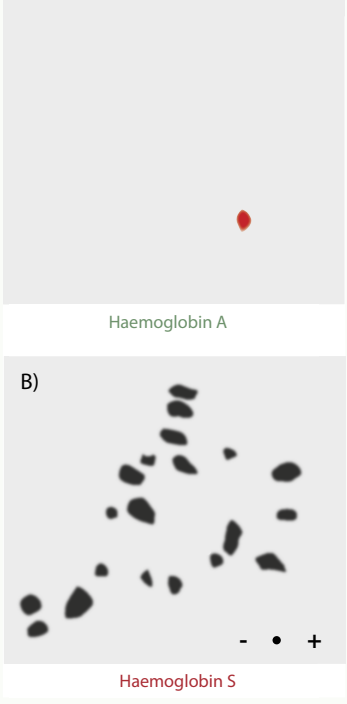


Figure 13.37 Peptide mapping to show the difference, highlighted in red, between (A) normal and (B) sickle cell haemoglobins.

Patients who are homozygous for SCA present with crises of intense pain that can occur anywhere in the body caused by blockage of the capillaries. Crises tend to occur when the circulation is slow or when there is hypoxia. Only about 15 seconds of low oxygen tension are required to produce sickling, so when the circulation is reasonably rapid there is insufficient time for this to happen. Clinical complications of SCA are highly variable and the clinical consequences may include acute chest syndrome, aplastic and bone pain crises, megaloblastic erythropoiesis, proneness to infections (especially by *Pneumococcus*, *Salmonella* and *Haemophilus* spp.) due to hyposplenism, and stroke. Acute chest syndrome is a common form of crisis in children with SCA and is sometimes fatal. It occurs in about 40% of all people with SCA. It is characterised by severe chest pain and difficulty in breathing. It is probably caused either by a chest infection or by blocked pulmonary capillaries resulting from a blood clot.

Diagnosing and treating sickle cell anaemia

A diagnosis of SCA/SCD is based on reduced numbers of erythrocytes and the characteristic shapes in their morphology (*Figure 13.38*) when examined by microscopy. When SCA is possible or suspected in a pregnancy, a prenatal diagnosis may be carried out using embryonic DNA obtained by chorionic villus sampling or amniocentesis using restriction fragment length polymorphisms (RFLPs). This term refers to differences in the base sequences of DNA at sites recognised by restriction endonucleases. If a mutation alters the sequence at such a site, then the lengths of DNA products produced when the DNA is hydrolysed by a given restriction enzyme will differ from the nonmutated site. These differences can be seen using gel electrophoresis to separate the fragments, which can be detected in the gel with a suitable probe. In SCA, the mutation changes one nucleotide base in *HBB*, forming a new sequence, which the restriction endonuclease *Mst*II cannot hydrolyse. Consequently, different sized fragments of DNA are produced from the sickle cell globin gene compared with the *HBB* DNA (*Figure 13.39*). The enzyme *Mst*II was one of the first to be applied for the prenatal diagnosis of SCA; however, other restriction endonucleases are now used similarly for its detection. If SCA is prenatally detected, the parents may then make an informed decision about whether to continue with the pregnancy.

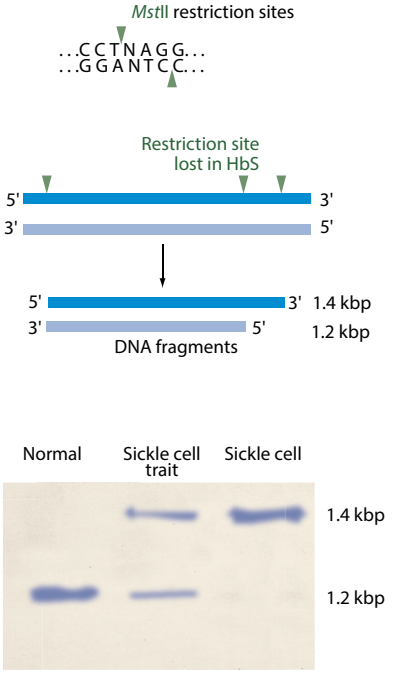


Figure 13.39 Outline of the use of restriction fragment length polymorphism with the restriction endonuclease *Mst*II in the prenatal diagnosis of sickle cell anaemia. Note the loss of the enzyme's target restriction site in the disease leading to the production of differently sized DNA fragments that can separated by gel electrophoresis for diagnostic purposes. HbS, sickle cell haemoglobin.

Treatment of SCA includes analgesics for the pain during crises and antibiotics, rest and vaccination against the likely life-threatening infections. The pain is often so acute as to require morphine. Sometimes inhalation of nitric oxide can help by producing a vasodilation, but this treatment only deals with the symptoms. Blood transfusions are also possible, but they can lead to iron overload, as well as other complications, as the transfused erythrocytes are removed from the circulation. Chelating agents, such as desferrioxamine or deferasirox (see *Figure 6.17*), may be used. It was observed that the severity of SCA in some populations was reduced by the presence of high concentrations of HbF, which is an excellent transporter of O_2 and, of course, does not cause sickling (*Section 13.3*). Butyrate and hydroxyurea (*Figure 13.40*) are used as therapeutic

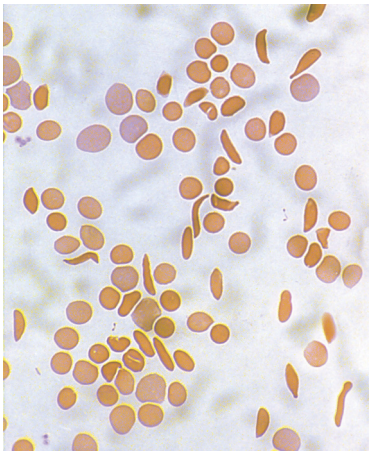


Figure 13.38 Photomicrograph of a blood smear from a patient with sickle cell anaemia. Note the sickled erythrocytes.

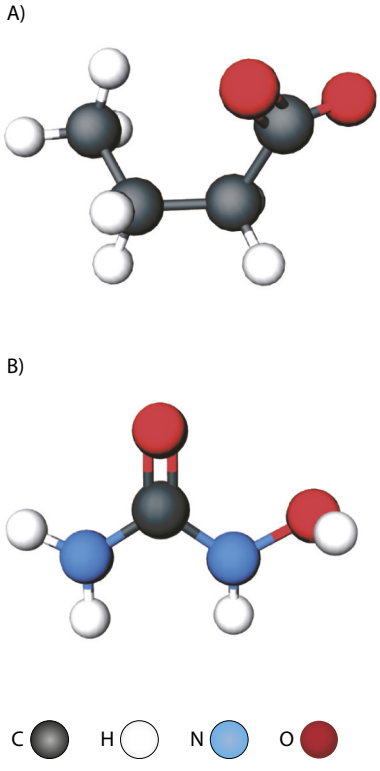


Figure 13.40 Computer drawn models of (A) butyrate and (B) hydroxyurea.

agents to try and induce a higher proportion of HbF in SCA patients. Butyrate seems to activate transcription of the γ globin gene so that HbF is produced in the adult. Both agents have met with reasonable success in treating patients and, in some cases, may be used synergistically to increase HbF by up to 20%, with a marked clinical improvement. Hydroxyurea is thought to selectively kill precursor cells in the bone marrow whilst sparing the erythroblasts that produce HbF. However, this compound is an antineoplastic agent (*Chapter 17*) and can result in liver damage in some patients. Haematopoietic stem cell (HSC) transplants (*Chapters 4* and *6*), of course, offer a permanent solution. Unfortunately, the lack of suitable donors limits the use of the treatment.

More recently, several molecular biological based therapies have become available. The humanised monoclonal antibody (see *Box 4.5*) crizanlizumab binds to P-selectins (see *Margin Note 5.5*) and blocks interactions between endothelial cells and erythrocytes (and also platelets and leukocytes), which prevents occlusion of capillaries by the sickled cells and eases crises pain. It is initially given as a single intravenous infusion dose, followed by a second two weeks later and then maintenance doses every four weeks. Some side effects are associated with the treatment, including arthralgia, diarrhoea, fever, gastrointestinal discomfort, myalgia, nausea and vomiting.

A limited number of patients with severe SCA have been treated by genetically modifying their HSCs collected by apheresis (*Chapters 4* and *6*). Two techniques have been used. In one, a lentivirus, modified to render it harmless, was used to transfer a therapeutic β globulin gene and its regulatory elements into the HSCs. The modified HSCs were returned to the patients. Following engraftment into the bone marrow, significant biosynthesis of native Hb occurred, with a marked improvement in patient symptoms and, no doubt, quality of life. In the second technique, harvested HSCs were transduced with a virus encoding a short hairpin RNA (shRNA, *Chapter 17*), which targets BCL11A mRNA. The transcription factor BCL11A plays a key role in the erythropoiesis transition from fetal to adult, that is, γ to β globin, biosynthesis (*Section 13.3*). Upon engraftment of the modified cells, the patients expressed HbF, which is an efficient O_2 carrier and not subject to sickling. Clinical symptoms of SCA were markedly reduced or absent. In essence, these patients resemble those with hereditary persistence of HbF (*Margin Note 13.2*).

The prognosis for SCA patients has improved, particularly over the last 20 or so years. However, the life expectancy of patients is variable, given that the severity of SCA differs between patients, that the onset of crises are unpredictable and that the treatment experienced varies. In the UK, childhood mortality is rare, with 99% of children surviving to adulthood; however, this is not the case in developing countries (*Box 13.9*). Unfortunately, even with appropriate clinical intervention, the survival rate is still poorer than the general population, with a mean age of death thought to be 67 years in the UK and 58 years in the USA.

BOX 13.9 SICKLE CELL TRAIT AND MALARIA ⏎

Sickle cell anaemia is present in approximately 8% of African Americans and may be as high as 45% in some African populations. Approximately 225 000 children are born with sickle cell anaemia in sub-Saharan Africa. Without intervention, 50–90% of affected children in many sub-Saharan African countries die before their fifth birthday. Heterozygotes, those with genes for haemoglobin A (HbA) and sickle cell haemoglobin (HbS), have sickle cell trait. Individuals typically have about 30% HbS, and life expectancy is about the same as for unaffected persons. The condition causes relatively few problems except at high altitudes and when flying in nonpressurised aircraft. It may be asked why such a deleterious gene should persist? The answer is probably that the possession of a single HbS gene, that is, sickle cell trait, increases resistance to malaria caused by *Plasmodium* spp. (*Chapter 2*). Malaria is typically endemic in areas where the sickle cell trait reaches high levels (*Figure 13.41*). The same may also be true of thalassaemia (see below). The reason for resistance seems to be that as *Plasmodium* parasites grow in the erythrocytes, they lower the intracellular pH and generate hydrogen peroxide. The lower pH promotes sickling of the erythrocytes, and the hydrogen peroxide damages cell membranes, which become more permeable to K^+. The resulting intracellular decrease in K^+ kills the parasites.

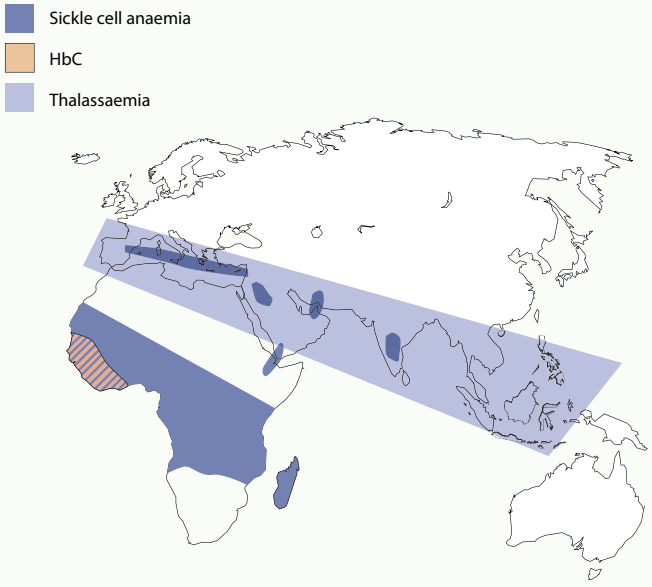


Figure 13.41 Areas of the world where sickle cell disease and thalassaemia are most common. HbC, haemoglobin C. ⏎

Thalassaemias

Thalassaemia (from *thalass*, the Greek for 'the sea') is a group of haemoglobinopathies. They were discovered originally in populations living near the Mediterranean Sea. However, thalassaemias are also relatively common in southeastern Asia, the Philippines and China. Thomas Cooley (1871–1945) first accurately described them clinically in 1925 and the disease was called 'Cooley's anaemia', a term now reserved for β thalassaemia major or transfusion-dependent β thalassaemia (see below).

In the late 1930s, thalassaemia was shown to be an inherited disorder. Thalassaemias seem to originate from point or very small deletion mutations; the more copies of the gene that are missing or inactive, the greater the severity of the disease. It was not until protein analytical techniques improved in the 1960s that the disease was shown to result from an imbalance in the relative amounts of α and β globins biosynthesised. Precipitation of the excess free globin within the erythrocytes, which are hypochromic and microcytic and have short circulating lives, produces the anaemia (*Figure 13.42*).

The kidneys secrete increased quantities of EPO (*Figure 13.28*) but, unfortunately, because erythroblast maturation is defective, this stimulates an ineffective erythropoiesis exacerbating the clinical presentation. Thus, the severe anaemias associated with some thalassaemias result in the bone marrow expanding in all parts of the skeleton, leading to skeletal deformities. If this occurs within the backbone and compresses the spinal cord it can cause intense pain. The erythrocytes have shortened lifespans and the spleen removes the abnormal erythrocytes leading to splenomegaly.

The several types of thalassaemia are a consequence of there being more than one globin gene, and a mutation may not affect them all. Consequently, there may not be a complete absence of a given globin and the disease will be less severe. In α thalassaemias there is a complete or partial failure to produce α globin(s), while in β thalassaemia the defect is a reduced or total failure to synthesise β globin(s).

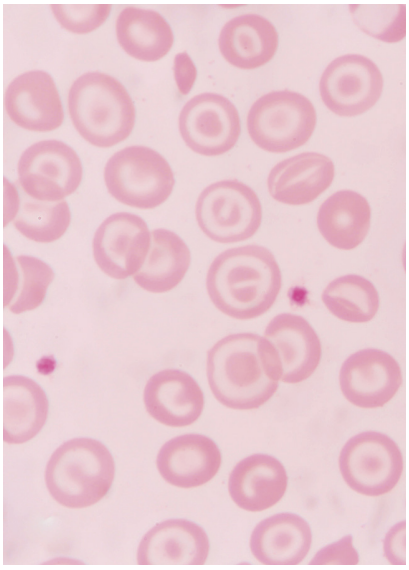


Figure 13.42 Photomicrograph of blood from a thalassaemia patient showing microcytic and hypochromic erythrocytes. Courtesy of Shutterstock.com/Chadsikan Tawanthaisong. ⏎

α *Thalassaemias*

Over 130 metabolic defects are associated with α thalassaemias. The condition can vary from one in which only a single α globin is missing, producing a 'silent' mutation that is, practically, a symptom-free, carrier state, to other forms where two, three or all four genes are absent or inactive (*Table 13.6*).

Syndrome	Genotype	Number of α genes	Clinical severity	Haemoglobin
Silent carrier	-α/αα	3	None	Functional
HbH disease	--/αα	2	Asymptomatic	Variable
	-α/α-	2	Asymptomatic	Variable
	--/-α	1	Severe microcytosis	About 25% HbA, mostly β_4
Bart hydrops fetalis syndrome	--/--	0	Lethal *in utero*	Mostly γ_4, little β_4

HbH, haemoglobin H.

Table 13.6 Characteristics of α thalassaemias. ⏎

The clinical severity of α thalassaemias generally correlates with the number of defective α globins present. Indeed, the presence of a single functional α globin gene is usually sufficient to preclude serious illness. Haemoglobin H (HbH) disease is characterised by a decreased biosynthesis of α globin leading to the formation of two abnormal Hb tetramers. Haemoglobin H precipitates in older erythrocytes and under oxidant stress, such as when the patient is treated for malaria with primaquine. However, the severity of the condition is variable; the genotypes --/αα and -α/α- have relatively mild clinical presentations, while --/-α results in a much more severe anaemia.

Some of the mutations that produce α thalassaemias are referred to as nondeletion mutations. For example, the mutation in the usual STOP codon of the α globin gene, UAA to CAA, means transcription continues to the next STOP codon. It appears that the extended mRNA is unstable and leads to a reduced rate of Hb synthesis, although its translation produces Hb Constant Spring, which has α globin molecules extended by 31 amino acid residues. The incidence of Hb Constant Spring is relatively common in

southeastern Asian populations. If Hb Constant Spring is present together with a lack of one of the α globin genes, then a severe HbH anaemia results. However, in heterozygotes, only about 1% of the α globin produced is the Constant Spring mutant type.

The absence of all α globin genes (genotype --/--) results in Bart hydrops fetalis syndrome, which is a lethal condition *in utero* or soon postpartum. Haemoglobin Barts is found in umbilical cord blood and arises because of a lack of α chains but normal production of the fetal γ globins; thus, Hb Barts has the composition γ_4. If the baby survives, due to intrauterine or immediate postpartum blood transfusions, β chain biosynthesis begins and β_4 (HbH) tetramers form. Unfortunately, neither γ_4 nor β_4 cooperatively binds O_2 and, indeed, they bind it so tightly it cannot be delivered to the tissues. Sadly, the condition is characterised by a variety of serious congenital clinical problems, and death occurs within five years.

β *Thalassaemias*

β Thalassaemia is the reduced biosynthesis of β globin with or without the production of smaller amounts of δ or γ globins. It results from more than 350 mutations in the β globin genes. In some ways, it would be more appropriate to call it *thalassaemias of the β globin gene family* because the genes for the single β globin and δ and γ globins are all grouped at the same locus, 16p13.3. Hence, β thalassaemia is not a single disease: the molecular defects in this and related disorders are highly heterogeneous. These include many single nucleotide substitutions affecting the expression of β globin genes, nonsense and frameshift mutations in the exons, point mutations at the intron–exon splice junctions (*Margin Note 13.7*), and mutations in the 5′ region and the 3′-polyadenylation sites (*Margin Note 13.10*).

Several heterozygous states are also known but these usually only give rise to mild clinical symptoms. In some forms of β thalassaemia, abnormal δ–β fusion polypeptides may be formed. For example, in Hb Lepore, there is a fusion gene caused by nonhomologous crossing over between the δ gene on one chromosome and the β gene on the other. Thus, normal β and δ genes are absent.

β Globin biosynthesis is required following birth (*Figure 13.13*) and β thalassaemic babies are usually born normally at term. Clinical problems begin two to six months later, when γ globin biosynthesis, and therefore the amount of HbF, has declined.

The three main types of β thalassaemia are:

1. β thalassaemia trait, also called β thalassaemia minor
2. β thalassaemia intermedia
3. β thalassaemia major, formerly called Cooley anaemia

Recently, the more clinically descriptive terms, nontransfusion-dependent thalassaemia for the intermedia type and transfusion-dependent rather than β thalassaemia major, have been widely introduced.

β Thalassaemia trait is the heterozygous form with only one mutated β thalassaemia gene. Often, the result is a patient who is asymptomatic with the mildest of anaemias. Both thalassaemia intermedia and major are homozygous conditions. The former generally presents later in life as a mild or moderate anaemia. In contrast, β thalassaemia major is a crippling disease of childhood, characterised by the precipitation of excess α globins, destruction of β thalassaemia erythrocytes in the bone marrow and circulation, and deficiency of functional Hb tetramers. The strongly depressed or absent β globin biosynthesis causes massive erythroid proliferation and skeletal deformities. The copious quantities of iron released from the rapidly destroyed erythrocytes are deposited in the heart, liver and pancreas, which become fibrous and dysfunctional, and in the skin, leading to a freckled and brownish appearance. Splenomegaly may be so pronounced as to impede respiration and cramp other abdominal organs.

Diagnosing and treating thalassaemias

A preliminary diagnosis of thalassaemia(s) can be obtained from the patient's general clinical symptoms, the anaemia and the presence of splenomegaly. This can be confirmed by a combination of laboratory tests, including measuring erythrocyte indices (*Table 13.3*), and by Hb and DNA analyses. Determining the number of erythrocytes and abnormalities in their sizes, shapes and colour is a primary screen for thalassaemia(s), where poikilocytes (the abnormally shaped products of the spleen when removing target erythrocytes) and hypochromic ones with a clear centre and darker rim containing precipitated Hb are observable (see *Figure 13.42*). However, microcytosis and decreased Hb content are characteristic of all thalassaemias, and this does not discriminate between thalassaemia trait and iron deficiency anaemia or between α and β thalassaemias. Both automated high-performance liquid chromatography (HPLC) and capillary zone electrophoresis (CE) furnish qualitative and quantitative data on the relative amounts of

Margin Note 13.10
3′-Polyadenylation ⏎

3′-Polyadenylation is a posttranscriptional modification of the nascent RNA. It consists of adding an extended sequence of adenine nucleotides to the 3′ end of RNA. This contributes to the stability of the mRNA formed and aids its transport from the nucleus to the cytoplasm.

α and β Hb subunits rapidly. Only DNA analyses can identify specific thalassaemia mutations, such as large deletions in α and point mutation in β thalassaemias. Various molecular biological techniques can be employed, including real-time polymerase chain reaction (PCR) and sequencing the patient's DNA using next-generation DNA sequencing methods (*Chapters 3* and *17*).

Patients with α thalassaemia trait, in general, do not require treatment. The condition HbH has a range of presentations. Relatively mild conditions can be treated with folic acid and iron supplements. In others, for example nondeletion types, such as Hb Constant Springs, repeated blood transfusions are necessary, and these must be accompanied by chelation therapy (see below) to remove excess iron. Splenectomy may be necessary. Prenatal diagnosis of Bart hydrops fetalis syndrome requires careful and considerate counselling to the parents, given the lack of effective care and the morbidity outcome for the baby.

In general, β thalassaemias are more severe, cause more patient suffering and are more expensive to treat because more medical intervention is required than is the case with α thalassaemias. Treatments resemble those for SCA. Rest, analgesics for the pain and folic acid supplements are required because of the haemolytic anaemia.

β Thalassaemia trait and intermedia patients are nontransfusion dependent, although they may require blood transfusions at specific times, for example in pregnancy or during acute infections. However, β thalassaemia major patients are transfusion dependent and, in the absence of curative treatment, require regular blood transfusions all their lives. Extensive transfusions must be combined with chelation therapy, with, for example, desferrioxamine, deferiprone or deferasirox (*Figure 6.17*) to remove excess iron. This is necessary because the body has no real excretion route for iron, and an iron overload may be fatal due to deposition cardiomyopathy by the second decade of life. Transfusions may also put the patient at risk of hepatitis and AIDS, especially in developing countries.

The hormone hepcidin (*Figure 13.43*), produced in the liver, regulates the expression of ferroportin, which helps regulate the absorption of iron from the gastrointestinal tract (GIT; *Chapter 11*), plasma iron concentrations and tissue iron distribution. In all severe anaemias, the secretion of hepcidin is suppressed and iron absorption increased, which can result in iron overload. Thus, administering hepcidin can help in the management of patients with β thalassaemia. Other treatments include giving hydroxyurea to augment HbF synthesis (see above) and luspatercept, a recently approved medication for treating adults with transfusion-dependent β thalassaemia. This is a recombinant fusion protein of the human activin receptor type IIb linked to the Fc portion of an IgG antibody (*Chapter 4*). The signalling pathway Smad2/3 is overactive in β thalassaemia patients, which slows down their maturation of erythrocytes. Luspatercept inhibits this pathway, therefore leading to an increase in erythrocyte production.

An HSC transplantation (*Chapter 6*) is a curative treatment for transfusion-dependent patients. Its drawbacks include a lack of suitable donors and the usual clinical risks associated with transplantation, such as graft versus host reactions. Recently, the lentivirus genetically modified with a therapeutic β globulin gene, that is, *HBB* and its regulatory elements, as described in the treatment of SCA (see above), has been approved as a treatment for adolescents and young adult patients suffering with transfusion-dependent β thalassaemia.

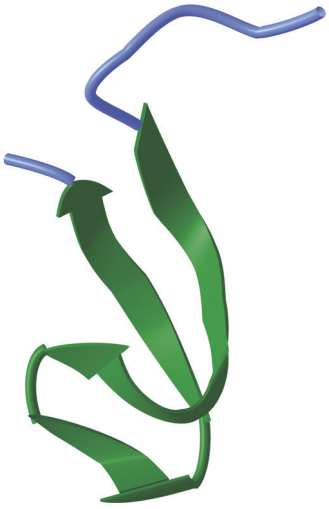


Figure 13.43 Molecular model of hepcidin. PDB file 2KEF. ⏎

13.8 GLUCOSE-6-PHOSPHATE DEHYDROGENASE DEFICIENCY

Erythrocytes lack mitochondria and cannot catalyse the TCA cycle but are able to generate ATP by glycolysis. They do possess a pentose phosphate pathway (PPP), sometimes called the hexose monophosphate shunt, which offers another route for the metabolism of glucose. Unlike glycolysis, however, the PPP does not form ATP but produces NADPH, a metabolic reducing agent essential for biosynthetic and some detoxification activities (*Margin Note 13.11*). The first reaction of the PPP is catalysed by glucose-6-phosphate dehydrogenase (G6PD). A deficiency of G6PD activity is one of the most common of what might be termed *enzymopathies*. The location of *G6PD* is Xq28; therefore, its deficiency is a sex-linked condition affecting males. It is, however, carried by females who have half the normal level of the enzyme activity (*Chapter 15*).

Genetic deficiencies of G6PD (*Figure 13.44*) are common in the Mediterranean region, Africa, the Middle East and South East Asia. It is estimated that about 500 million people are affected, making G6PD deficiency the most common inherited disease, although many are largely asymptomatic. However, any of them could develop an acute, indeed severe, haemolytic anaemia in response to environmental cues. Approximately 230 different mutations of *G6PD* have been reported. Almost all are missense types and

Margin Note 13.11 ⏎
Why NADH or NADPH?

The coenzymes NADH and NADPH are both reducing agents. However, NADH is largely used by mitochondria in electron transport to drive the formation of ATP. This contrasts with the mainly anabolic roles of NADPH in reductive biosynthetic reactions and as a coenzyme for cytochrome P-450s in the detoxification of xenobiotics (*Chapter 12*).

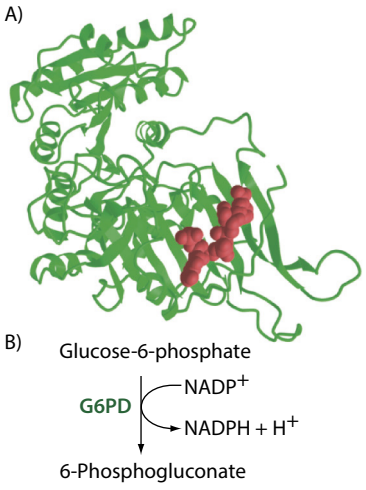


Figure 13.44 (A) Molecular model of glucose-6-phosphate dehydrogenase (G6PD). Its bound $NADP^+$ coenzyme is highlighted. PDB file 6E08. (B) The reaction catalysed by G6PD.

approximately 50% are sporadic. Mutations that lead to an unstable enzyme or one with abnormal kinetics result in reduced enzyme activity, and the severity of the haemolytic anaemia is proportional to the decrease in enzyme activity compared with reference values. The deficiency mostly affects erythrocytes because they have a long life in circulation and cannot perform protein synthesis to replace the defective enzyme. The most prevalent mutation is the African or A type, present in about 11% of Black populations. The degree of deficiency is mild, with enzyme activity being about 10% of usual levels, and erythrocytes can manufacture sufficient NADPH under normal circumstances.

In erythrocytes, the requirement for NADPH relates largely to its role as an essential source of reducing power for cytochrome b_5 reductase (methaemoglobin reductase), and the need to replenish reduced glutathione (GSH). The activity of NADP–cytochrome b_5 reductase in maintaining the Fe^{2+} oxidation state in Hb was described in *Section 13.5*. A constant supply of NADPH is required from the PPP to prevent the precipitation of denatured Hb and Heinz body formation and maintain erythrocyte integrity. Glutathione is a γ-linked tripeptide of glutamate, cysteine and glycine residues containing a thiol (-SH) and is also discussed in *Chapters 12* and *18*. Its major role in erythrocytes is eliminating hydrogen peroxide (H_2O_2) a toxic by-product of methaemoglobin formation and superoxide dismutase activity. Hydrogen peroxide can react with unsaturated fatty acid residues in the erythrocyte membrane forming hydroperoxides, damaging the membrane and causing premature cell lysis (*Chapter 10*). The hydroperoxides are removed by the activity of glutathione peroxidase, which requires GSH as a reducing agent:

$$\text{R-O-O-H} + 2\,\text{GSH} \xrightarrow{\text{Glutathione peroxidase}} \text{R-OH} + \text{GSSG} + \text{H}_2\text{O}$$

peroxide; reduced glutathione; oxidised glutathione

The limited amount of glutathione in the cell means the reduced form, GSH, must be regenerated in an NADPH-requiring reaction catalysed by glutathione reductase. Note that the NADPH is used to reduce the enzyme bound coenzyme, FAD, which, in turn, reduces the oxidised glutathione (GSSG; *Figure 13.45*).

Most patients can make sufficient NADPH under normal conditions, and the defect only becomes apparent when the person, for example, takes a drug, such as the antimalarial primaquine; inadvertently ingests a chemical like naphthalene (mothballs), which greatly increases the demand for NADPH; is infected with a virus or bacterium; or eats broad beans (*Vicia fava*; see *Box 13.10*). Many types of drugs, besides antimalarials, require NADPH for their detoxification (see *Table 13.7*). In individuals with a severe form of the disease, the oxidative stress can precipitate a clinical crisis with severe haemolytic anaemia and a loss of 30–50% of the erythrocytes. Heinz bodies (see *Margin Note 13.8*) may be present. The urine may turn black because of the high concentrations of Hb and its degradation products, and a high urine flow must be maintained to prevent renal damage.

Diagnosis and treatment of glucose-6-phosphate dehydrogenase deficiency

People with G6PD deficiency generally do not have symptoms unless they are exposed to drugs and chemicals (see *Table 3.10*) or a viral or bacterial infection that produces

Margin Note 13.12 Glucose-6-phosphate dehydrogenase deficiency and malaria

Female carriers of G6PD deficiency, like SCA patients, are more resistant to the malarial parasite. This resembles the link between SCA and malaria. Indeed, the two diseases have similar geographical links with malarial areas (*Figure 13.41*). There is evidence that G6PD-deficient erythrocytes infected with *Plasmodium falciparum* (*Chapters 2* and *3*) can be detected by macrophages at an early stage in the development of the parasite and eliminated.

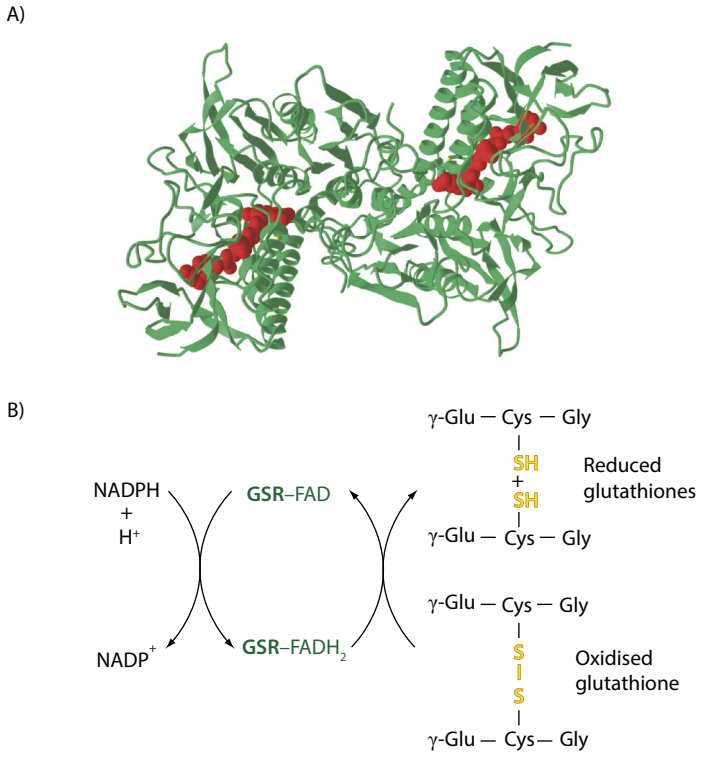


Figure 13.45 (A) Molecular model of glutathione reductase (GSR), with its bound FAD coenzymes highlighted. PDB file 3GRS. (B) The reaction catalysed by GSR.

Group	Examples
Antimalarial drugs	Pamaquine, primaquine, pentaquine, atabrine, quinine[a]
Analgesics	Aspirin (high dose), phenacetin
Antibacterial drugs	Chloramphenicol[a], nitrofurantoin, furazolidine
Sulphonamides	Sulphacetamide, sulphanilamide, sulphapyridine
Sulphones	Dapsone, thiazosulphone, diphenylsulphone
Arsenicals	Neoarsphenamine
Chemicals	Methylthioninium chloride (methylene blue), naphthalene, phenylhydrazine, toluidine blue, nitrite, ascorbic acid (large doses), nalidixic acid

[a] Only in the Mediterranean type.

Table 13.7 Examples of drugs and chemicals that can cause acute haemolytic anaemia in individuals with glucose-6-phosphate dehydrogenase deficiency.

an oxidative stress, as described above. The most common associated medical problem is haemolytic anaemia, which presents as fatigue and paleness or yellowing of the skin and whites of the eyes (jaundice), dark urine and an increased heart rate. Splenomegaly may also be present. However, many underlying conditions can cause haemolytic anaemia, including some other genetic diseases, and autoimmune forms can occur. If on initial examination, a patient is suspected of G6PD deficiency, then fluorescent spot or dye reduction screening tests may be used to asses G6PD activity. In the spot test, the NADPH produced converts a reagent to a product that fluoresces when viewed using ultraviolet light. Fluorescence must be produced within 10 minutes to eliminate a deficiency. The intensity of the fluorescence is a measure of G6PD activity; hence, the test is only semi-quantitative. Those tests that produce a change in absorbance when a dye is reduced, however, give a quantitative value of the enzyme activity. Note that G6PD activity is highest in immature erythrocytes or reticulocytes and declines as the mature

erythrocytes age. During an attack of haemolytic anaemia, the older erythrocytes are destroyed first, leaving behind reticulocytes and so giving a higher reticulocyte count. Consequently, measuring G6PD activity in erythrocytes following a crisis can lead to spuriously high values. Molecular biological procedures can be considered in cases where the DNA lesion needs to be identified. Methods based on the PCR are available, while advanced next-generation sequencing methods allow all possible erythrocyte enzyme deficiencies to be investigated simultaneously. Identifying the carrier status of a woman who is possibly heterozygous for G6PD deficiency is only possible using molecular biological methods.

The treatment of haemolytic anaemia induced in G6PD deficiency involves addressing the underlying cause of the condition, removing the agents that induced the attack or treating the infection. Rest, analgesics and fluids are appropriate for milder cases. Blood transfusions to alleviate severe anaemia are necessary. If the patient survives the crisis, with or without transfusions, recovery will usually occur quite rapidly, as the reticulocyte count increases even if the individual continues to take the noxious agent, because reticulocytes can biosynthesise G6PD.

BOX 13.10 FAVISM

Favism is an attack of haemolytic anaemia brought on by the consumption of fava or broad beans or perhaps even their pollen (*Vicia fava*; *Figure 13.46*) in some, glucose-6-phosphate dehydrogenase (G6PD)-deficient individuals. Infants are especially susceptible. It is not understood why only some patients are susceptible, but it has been suggested that another mutation must also be present for favism to be expressed. Also, broad beans contain small quantities of toxic glycosides, which, like antimalarial drugs, increase the demand for NADPH, but this does not explain the effects of pollen. Broad beans are a popular item in the diet in Mediterranean countries, although they were the only edible bean known in Europe until different types were introduced from the New World. Favism is frequently fatal unless a large volume of blood is transfused promptly. The philosopher Pythagoras (ca 570–495 BCE) forbade his followers from eating broad beans, not on the grounds of possible NADP deficiency, but because he believed they contained the souls of the dead.

Figure 13.46 Broad bean plant (*Vicia fava*).

13.9 CLINICAL ASPECTS OF CLOTTING

Clinical problems occur if the blood fails to clot, for example in vitamin K deficiency (*Box 13.3*) and haemophilia (see below), or if it clots too easily or inappropriately, as in a coronary thrombosis (*Chapter 14*). Thus, analytical methods to clinically investigate the clotting abilities of blood samples are necessary. Blood samples require the addition of Ca^{2+} chelating agents, such as oxalate, citrate, EDTA (see *Figure 12.28*) or heparin (*Figure 13.24*), to prevent them clotting and centrifugation to remove blood cells. The addition of excess Ca^{2+} or the removal of the chelating agents to the remaining plasma allows it to clot. Thus, the clotting of blood from patients can be investigated in the pathology laboratory and any defects identified to help with the diagnosis and monitoring of treatments. For example, it is also possible to determine the activity of the intrinsic pathway by estimating the partial thromboplastin time. Calcium ions and phospholipids are added to plasma, which is then exposed on a surface to activate factor XII and the time taken for a clot to form is noted. The activity of the extrinsic pathway can be estimated by

determining the prothrombin time. Thromboplastin is added to blood from the patient into a tube containing citrate to chelate the Ca^{2+} and the time taken for clotting to occur is noted. The thrombin time is a measure of the activity of thrombin and is found by adding thrombin to the plasma sample and timing the interval for a clot to form. Some patients with haemophilia A (see below) may have normal prothrombin times because their concentration of TF is high.

The treatment and monitoring of clotting disorders are specific to the clinical problem; here, only four will be discussed:

1. von Willebrand disease
2. haemophilias
3. thrombocytopenias
4. thrombocytosis

von Willebrand disease

von Willebrand disease (vWD) is the most common inherited bleeding disorder, affecting an estimated 90 000 people globally, although this title has also been attributed to several clotting conditions. It arises from mutations in *VWF* located at 12p13.3. It is inherited as an autosomal dominant condition (*Chapter 15*). The gene encodes factor X or von Willebrand factor (vWF)/protein. This is a large, indeed the largest, plasma protein. It is multimeric, consisting of a variable number of subunits giving it a range of M_r of $500 \times 10^3 - 20 \times 10^6$. von Willebrand factor is a carrier of clotting factor VIII in the plasma and has metabolic functions in platelet adhesion.

The platelet dysfunction in vWD results mainly in bleeding of mucous membranes, for example spontaneous epistaxis, gingival blood loss and menorrhagia (prolonged menstrual cycles), and prolonged bleeding after delivery, trauma or surgery. The condition is divided into three main subtypes depending upon the mutations involved. These are named, with no great originality, vWD types 1, 2 and 3. The first accounts for 60–80% of all cases. It is distinguished by a partial deficiency of circulating vWF. Type 2 constitutes 10–30% of cases, with relatively minor abnormal forms of vWF. Type 3 forms 1–5% of all cases but is characterised by a severe deficiency of vWF in plasma, with less than 1% of normal concentrations.

Diagnosing and treating von Willebrand disease

A definitive diagnosis of vWD and its subtypes is difficult. Care must be taken to distinguish it from haemophilia. For example, factor VIII clotting activity may be assessed to determine if there are abnormally low levels of factor VIII activity (see below). Patients undergo evaluations that include their past medical history and checks for bruises or signs of recent bleeding. Several blood tests are available, including determining vWF antigen, which is a measure of the concentration of the factor in blood, or measuring its activity to determine how well it functions during clotting. The presence of multimers of vWF in blood can be investigated to assess which protein complexes are present and how they are degraded, which may help identify the type of vWD present. If vWD is diagnosed in a patient, then other family members should be tested to determine if the disease runs in the family.

Treatment of vWD depends on the type and severity of the condition and any previous response(s) to therapy. Desmopressin (*Figure 13.47*), an analogue of vasopressin (*Chapters 7* and *8*), can be used to treat mild to moderate cases of vWD, given that it

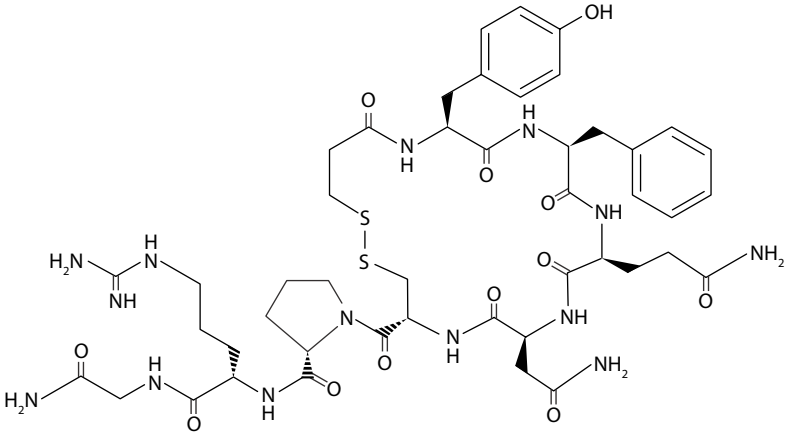


Figure 13.47 Desmopressin. Compare to *Figure 7.20*. ⏎

stimulates the release of vWF from endothelial cells. The replacement of vWF using infusions containing concentrated vWF and factor VIII is an option in cases of severe bleeding. Clot stabilisers, for example antifibrinolytic medications, that can stop bleeding by inhibiting the dissolution of blood clots may also be of use.

Haemophilias

Haemophilias, the 'bleeding diseases', are caused by mutations that produce defective or a lack of clotting factors. The bleeding times are normal in these disorders but the clotting (or coagulation) times are prolonged. The two most common are haemophilias A and B.

Haemophilia A

Haemophilia A, in which factor VIII (*Figure 13.48*) is deficient, is the most common of the haemophilias. Its incidence varies from 1 in 5000–10 000 men. The deficiency of factor VIII is caused by mutations in its recessive gene at Xq28 (*Margin Note 13.13*). The presence of a defective allele in males on their sole X chromosome (X^hY) means they will present with the disease. Female carriers have a recessive gene on one X chromosome and a normal gene on the other (X^HX^h): their blood clots normally. Given that haemophilia affects only about 0.02%, or less, of males in the population, the frequency of X chromosomes carrying the mutated allele is about 0.0002 among human males. Therefore, among human females, the occurrence cannot exceed $(0.0002)^2$, so female haemophiliacs (X^hX^h) are very uncommon. Thus, haemophilia affects mainly males and is normally transmitted by a female carrier who shows no symptoms.

Margin Note 13.13 ⏎ The gene for factor VIII

The human gene for factor VIII, *F8*, is sited at Xq28. It was cloned as long ago as 1984. It is enormous, consisting of 186 kilobases and forms about 0.1% of the DNA in the X chromosome. It is subjected to various genetic defects, including deletions, point mutations and insertions. Spontaneous mutations in the factor VIII gene are relatively common.

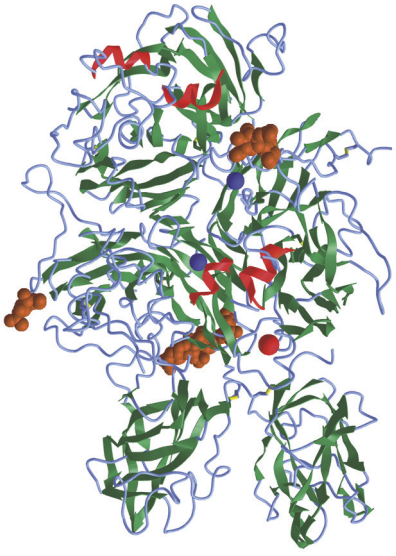


Figure 13.48 Molecular model of factor VIII. Attached glycans are shown in brown, while Ca and Cu ions are represented as red and blue spheres, respectively. PDB file 3CDZ. ⏎

The main symptom of haemophilia A is a tendency for uncontrolled haemorrhages, that is, blood loss, either spontaneous or from even small injuries. If untreated, internal bleeding into the joints and muscles may occur and there is a risk during surgical operations.

Haemophilia A is generally diagnosed in early childhood following observations that minor cuts bleed extensively and the child easily bruises and develops haematomas. An initial diagnosis can be confirmed by determining the plasma concentrations of factor VIII, which in patients with severe haemophilia A is less than 5% of normal.

Treatment involves educating patients, and the parents of affected children, regarding the condition. Minor cuts can be managed with gauze and pressure bandages. In the past, haemophilia A was treated with blood transfusions. This was replaced by using factor VIII purified from donated bloods. Advances in haemophilia A treatment involve infusions of factor VIII prepared by recombinant DNA technology. This eliminates any chance of infection but is expensive. Factor VIII has a plasma half-life of 12 hours, so it needs to be administered twice daily to maintain the required therapeutic concentrations in the plasma. Patients should be vaccinated against hepatitis A and B. Additional clinical advances include home care treatment and immunosuppressant therapy. All have significantly improved the quality of life for haemophiliac A patients and increased life expectancy to within 10 years of that of males in the general population.

Haemophilia B

Haemophilia B, in which factor IX (Christmas factor; *Figure 13.49*) is deficient, is approximately seven times less common than haemophilia A, with an incidence of about 1 in 30 000–40 000 male births in USA but is the second most common form of haemophilia. It is also called Christmas disease because the factor was first found to be missing in a patient, Stephen Christmas, in 1952.

The deficiency of factor IX is caused by mutations in its recessive gene at Xq27. The presence of a defective allele means haemophilia B resembles the A form in being inherited as an X-linked recessive disorder; males are affected and females are carriers. Symptoms vary in severity depending upon the extent of factor IX deficiency. The treatment and management of haemophilia B resemble those of A; patients inject themselves, usually on a weekly basis, with factor IX. However, there is considerable interest in a new type of adeno-associated virus (AAV) gene therapy, called FLT180a, to treat severe and moderately severe haemophilia B patients. The therapy delivers a functional copy of the gene directly to patient tissues to compensate for the defective copy. Thus, when successful, newly made factor IX is released into the blood. In a trial reported in 2023, 9 out of 10 patients received a single treatment with FLT180a that led to the sustained production of factor IX by the liver, removing the need for regular replacement therapy.

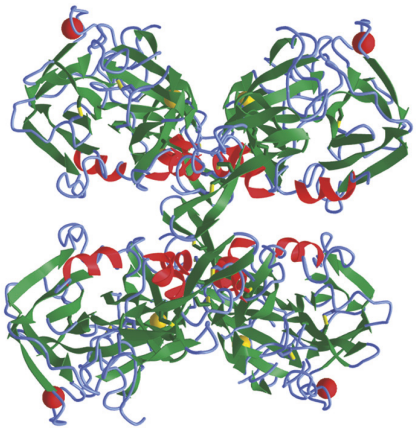


Figure 13.49 Molecular model of factor IXa/thromboplastin/Christmas factor. Ca^{2+} and Na^{+} are represented as red and yellow spheres, respectively. PDB file 6MV4.

BOX 13.11 HAEMOPHILIA AND QUEEN VICTORIA

Haemophilia is associated with numerous members of the royal families of Europe. Queen Victoria (*Figure 13.50*) appears to have received a mutant allele from one of her parents and, as shown in *Figure 13.51*, she passed on the gene to subsequent generations. Her consort, Prince Albert could not have been responsible because male-to-male inheritance is not possible. One son of Victoria, Leopold, Duke of Albany, died of haemophilia at the age of 31 and at least two of her daughters were carriers. Various intermarriages allowed the disease or trait to spread from throne to throne across Europe; sufferers included the son of the last Tsar of Russia.

Figure 13.50 Queen Victoria.

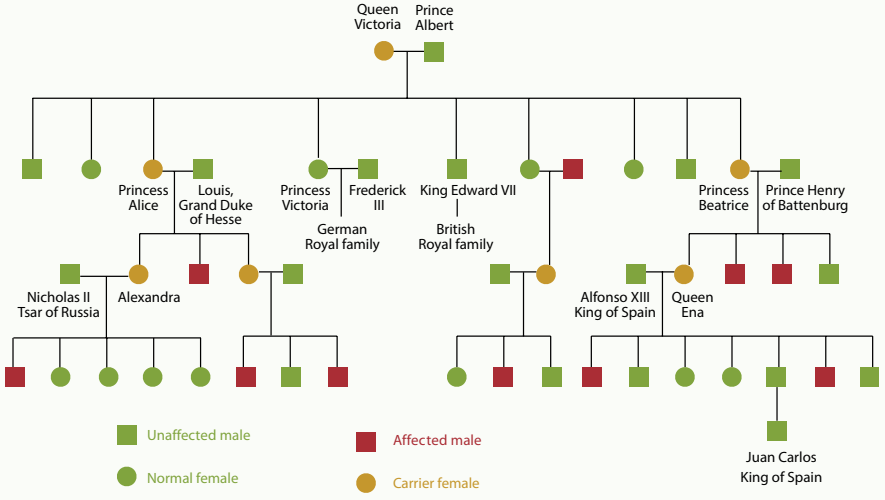


Figure 13.51 Partial family tree showing the inheritance of the haemophilia gene in descendants of Queen Victoria.

Margin Note 13.14 Haemophilia C

Haemophilia C, also called plasma thromboplastin antecedent (PTA) deficiency or Rosenthal syndrome, is caused by a deficiency factor XI. It affects 1 in 100 000, but is much more common in Ashkenazi Jews because of consanguinity. Haemophilia C is a milder form of haemophilia than A or B and affects both sexes.

Thrombocytopenia

Thrombocytopenia is a reduction in the number of platelets in the blood resulting in prolonged clotting times and the formation of clots that are soft and do not retract. It is the most common bleeding disorder. The condition can result from an insufficient number

of platelets produced, an excessive rate of their destruction or a combination of both. Several clinical conditions reduce erythrocyte formation in the bone marrow. Examples include aplastic anaemia, some leukaemias, viral infections such as cytomegalovirus, Epstein–Barr virus (glandular fever), Rubella (German measles), *Paramyxovirus* (mumps) and chicken pox caused by *Varicella zoster* (*Chapters 2* and *3*). Toxic chemicals (*Chapter 12*), arsenic and benzene for example, can also cause the condition. The spleen normally contains approximately 30% of the body's platelets. Thus, an enlarged spleen arising from some cancers (*Chapter 17*) or liver cirrhosis *(Chapter 12*) can lead to an increased entrapment of platelets and their catabolism. Iatrogenic thrombocytopenia may be caused by treatment with some drugs that reduce platelet number, for example certain diuretics and the antibiotic chloramphenicol, while other medications, such as quinine, rifampin and vancomycin, may lead to increased platelet destruction.

Thrombocytopenic purpura (TPP) is characterised by bleeding in the skin, producing purpura (brown and red bruises) or **petechiae** (small red to purple dotted areas; *Chapter 6*). The condition is often considered idiopathic (ITP) although in most cases antibodies are present that reduce platelet counts (autoimmune thrombocytopenia). Thrombocytopenic purpura may be acute or chronic. The former occurs in children often following a viral infection, as described above, and lasts weeks to months, normally with no lasting effects. The chronic form afflicts adolescents and adults and, unfortunately, is rarely self-limiting. Appropriate clinical management is associated with a mortality of approximately 1%. Thrombocytopenia presents in 5% of pregnant women as delivery approaches; however, the reasons for this are obscure.

The first symptoms of thrombocytopenia include bleeding gums and nose (epistaxis), skin that easily bruises, and bleeds giving pinpoint lesions. It can be a life-threatening condition and severe cases may be fatal from heavy bleeding in the brain or GIT.

Diagnosing and treating thrombocytopenia

Suspected thrombocytopenia should be investigated by a careful clinical history and physical examination. Clinical tests involve determining platelet counts, assessing the size of the spleen and bone marrow biopsy.

If the causes are external agents and/or iatrogenic then the offending agent(s) need(s) to be eliminated immediately. Mild cases often improve without the need for further management. Severe cases require platelet transfusions, especially when active bleeding is occurring or the risk of it is high. Medications include corticosteroids, immunoglobulins, vitamin K supplements and the thrombopoietin receptor agonists (TPO-RAs) eltrombopag and romiplostim (*Figure 13.52*), which bind to the receptor and stimulate platelet production. Splenectomy is an option, particularly if the underlying problem is autoimmune in origin.

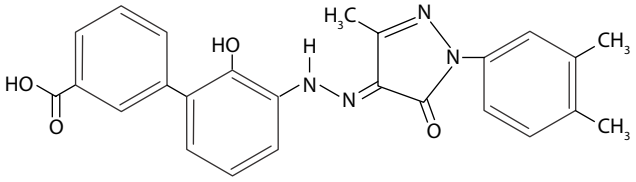


Figure 13.52 Eltrombopag.

Thrombocytosis

Thrombocytosis, in contrast with thrombocytopenia, is an abnormally increased number of blood platelets. Platelet numbers may exceed 450 000 /μL and increase the chances of thromboses. The condition can be primary, a myeloproliferative disease. Secondary thrombocytopenia may be caused by splenectomy or some other operations, or by inflammatory diseases or even as a response to extreme exercise. The condition is often asymptomatic and only discovered when blood tests are performed for another reason. When symptoms do, rarely, present, they may include an enlarged spleen, skin bruising of unknown cause, bleeding from places like the nose and gums, or blood in stools. Clinical complications associated with thrombocytosis are thromboembolisms that can occur in any of many organs, for example the brain giving stroke-like symptoms, the lungs leading to chest pains and shortness of breath, or the heart causing an attack (*Chapter 14*). In pregnant women, clinical complications include reduced fetal growth, abruption of the placenta or miscarriages. The usual treatments are to give antiplatelet (platelet-lowering) drugs, such as aspirin (see *Figure 12.13*), anagrelide (*Figure 13.53*), hydroxyurea (*Figure 13.40*) or interferon α (*Chapter 4*).

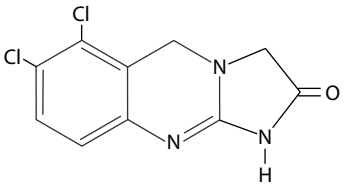


Figure 13.53 Anagrelide.

13.10 CASE STUDIES

CASE STUDY 13.1

Daniel, a four-year-old child of Jamaican origin, was brought into the hospital because of frequent severe headaches and abdominal pain. He had also been tired and sleepy for several months. His sclerae were yellow and abdomen was distended. Palpation revealed an enlarged spleen. His blood samples gave the following data (references ranges in parentheses):

Erythrocyte count	2.2×10^{12}	($4.4–5.9\times10^{12}$/L [men])
Haemoglobin	47	(140–180 g/L)
Serum bilirubin (unconjugated)	+++	(±)

A fresh smear of blood showed a few crescent-shaped cells and so haemolysed erythrocytes from Daniel and both his parents were investigated by electrophoresis (*Figure 13.54*).

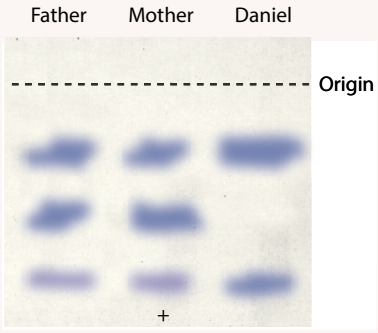


Figure 13.54 Electrophoresis of haemolysed erythrocytes from Daniel and his parents. ⏎

Questions

(a) What is the most likely diagnosis?
(b) Outline the sort of treatment that would be appropriate.
(c) What is the chance of the parents' next child having this condition?

CASE STUDY 13.2

Bill, a 57-year-old man, lost weight and complained of weakness, shortness of breath and a sore tongue and had difficulties with swallowing. On referral it was found that his skin was yellowish and his tongue was shiny and smooth. He said that he had gradually lost his appetite and recently had only taken liquid foods in order to avoid abdominal pain. He also mentioned numbness and tingling of the hands. His temperature was a little above normal at 39°C. Blood index measurements gave the following data (references ranges in parentheses):

Erythrocyte count	3.7×10^{12}	($4.4–5.9\times10^{12}$/L)
Total haemoglobin	82	(135 g/L)
PCV	37%	(40–52%)

A blood smear showed large erythrocytes of unusual shape. His bone marrow sample contained basophilic megaloblasts. Gastric secretion was small in volume, 0.3 L in 24 hours compared with the normal 2.5 L, and contained little hydrochloric acid.

Questions

(a) Calculate the values of MCV, MCH and MCHC.
(b) What is the most likely diagnosis and its most probable cause?
(c) Suggest plausible reasons for the gastric problems.
(d) How should Bill be treated?

CASE STUDY 13.3

Alan has SCA and has frequent sickling crises. His brother, Michael, is known to have sickle cell trait. Their sister, Liza, seems to be normal and does not show any symptoms. However, she is about to marry and wants to start a family. She is therefore concerned to know her sickle cell status. A Hb electrophoresis on her blood was carried out and showed that she had 53% HbA, 46% HbS and 1% HbF.

Questions

(a) What molecular biological test could be carried out to confirm Liza's Hb distribution (you may wish to consult *Chapters 4* and *17*)?
(b) If Liza's proposed marriage partner was heterozygous for SCA, what are the chances that their first child would have SCA?

CASE STUDY 13.4

Matthew, a nine-month-old male infant, was brought to the Accident and Emergency Department with a painful, expanding mass in his left thigh. His mother noticed this a few hours after he had fallen on a hard floor, and the child was in considerable pain. An X-ray showed there were no fractures, but the soft swelling was a haematoma caused by bleeding into the tissues. On questioning, Matthew's mother said that soon after he began to crawl Matthew's knees became painful and swollen. A coagulation disorder was suspected. Tests showed that Matthew had only about 5% of the normal concentration of plasma factor VIII.

Questions

(a) What is the most likely diagnosis?
(b) How should Matthew be treated?

CASE STUDY 13.5

Ruby visited her doctor complaining of extreme tiredness and pain in her lower spine. Her doctor noted her pallor, paleness beneath the eyelids and some slight deformities of her skull and facial bones. Blood tests indicated severe anaemia and a computed tomography (CT) scan of the bones of her spine showed an area of spinal cord compression in the upper lumbar region. These and additional clinical tests indicated a thalassaemia.

Questions

(a) Discuss Ruby's clinical presentation.
(b) Comment on possible treatments.

13.11 SUMMARY

The blood system has many components: erythrocytes, leukocytes and platelets suspended in the complex electrolyte- and protein-rich plasma. Haemoglobin transports oxygen from the lungs to other tissues. Anaemia results when Hb concentrations decrease and this is usually accompanied by reduced numbers of erythrocytes. Anaemias may be associated with normal-sized erythrocytes or ones that are microcytic or macrocytic, depending upon the underlying cause.

Many hundreds of mutations are known that affect Hb giving rise to haemoglobinopathies. These range from relatively mild conditions to ones that are severe and life threatening. Sickle cell anaemia is caused by a point mutation in the gene for β globin. This results in molecules of deoxyhaemoglobin aggregating to form fibrils, distorting the erythrocytes and making them less flexible when passing through capillaries. This results in a lack of oxygen in the tissues, severe pain and haemolysis. Thalassaemias are a group of relatively common haemoglobinopathies, caused by the defective biosynthesis of the α or β globins. Thalassaemias, again, vary from the relatively mild to very severe, depending upon the precise genetic defect. The most common blood disease is G6PD deficiency. Typically this is asymptomatic until the patient takes a particular drug that produces an excessive demand for glutathione in the erythrocytes and this leads to a severe haemolytic anaemia. Many defects of the blood clotting system are known. von Willebrand disease and haemophilias are clinical conditions associated with dysfunctional clotting of the blood. If not managed appropriately, these can be fatal disorders.

QUESTIONS

1. Which of the following statement(s) about iron-deficiency anaemia is/are TRUE?
(a) The erythrocytes are small and hypochromic.
(b) The serum iron-binding capacity falls.
(c) There is angular stomatitis and brittle hair.
(d) Deposits of haemosiderin will be found in the liver and spleen.
(e) The erythrocytes are microcytic in appearance.

2. Identify the leukocytes A and B in *Figure 13.55*.

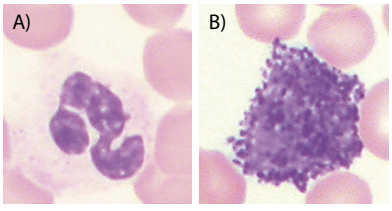


Figure 13.55 Two leukocytes.

3. *Figure 13.56* shows the electrophoresis of Hbs from four patients, A–D. In SCD, the change at position 6 in the β globin chain is from Glu to Val and in HbC is from Glu to Lys. Individual C is normal, with HbA present. What can be said about individuals A, B and D?

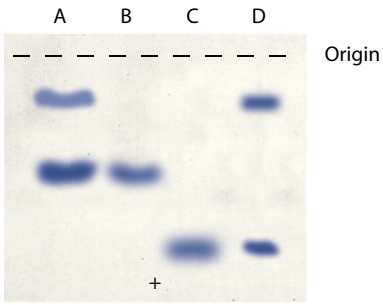


Figure 13.56 Electrophoresis of haemoglobins from patients A–D.

4. Explain why hereditary persistence of HbF does not have serious consequences.

5. Indicate in the following table whether the haematological indices listed are not affected (NA), decreased (D) or increased (I) in iron deficiency and pernicious anaemias.

Index	Iron deficiency anaemia	Pernicious anaemia
PCV		
MCV		
MCH		
MCHC		

6. Discuss the role of platelets in blood clotting.

7. Which of the following is the odd ONE out?
(a) Dapsone.
(b) Haemoglobin M-Saskatoon.
(c) Fe^{2+}.
(d) NADH cytochrome b_5 reductase.
(e) Methylthioninium chloride (methylene blue).

8. Which of the following statements are TRUE and which are FALSE?
(a) If blood is allowed to clot, the clear straw-coloured liquid remaining after removal of the clot is plasma.
(b) The formed elements of blood originate from totipotent stem cells.
(c) Approximately 2×10^{11} erythrocytes are biosynthesised daily.
(d) Iron is transported in the plasma by ferritin and stored in the liver and spleen cells in transferrin.
(e) The three major types of leukocyte are polymorphonucleocytes, lymphocytes and monocytes.
(f) The reduction of methaemoglobin converts the $Fe(II^{2+})$ in Hb to oxidised $Fe(III^{3+})$.
(g) Embryonic and fetal Hbs have greater affinities for dioxygen than adult Hb.
(h) Factor IXa is common to the intrinsic, extrinsic and final pathways of the clotting cascade.
(i) α and β antithrombin IIIs are natural clotting inhibitors; however, α antithrombin III has less affinity for heparin the the β form.
(j) Glutathione reductase contains the enzyme-bound coenzyme FAD.

9. Arrange the following two columns into their most appropriate pairings.

Normocytic anaemia	Petechiae
Pernicious anaemia	Glutathione
Megaloblastic anaemia	Mucous membranes
Precipitated Hb	Insufficient in number of erythrocytes
Crizanlizumab	Enlarged oval erythrocytes
β Globin gene family	Heinz body
Oxidised glutathione	NADPH
GSSG	16p13.3
vWD	P-selectins
Thrombocytopenic purpura	Macrocytic anaemia disease

FURTHER READING

Algrain L (2021) Beginner's guide to next-generation sequencing. *Biochemist (Lond)* **43**(6):58–64.

Ammouri W *et al*. (2020) Pernicious anaemia: mechanisms, diagnosis and management. *EMJ Hematol* **1**:71–80.

Angastiniotis M and Lobitz S (2019) Thalassemias: an overview. *Int J Neonatal Screen* **5**:16. doi: 10.3390/ijns5010016.

Babkir M *et al*. (2020) Iron-induced thrombocytopenia: a mini-review of the literature and suggested mechanisms. *Cureus* **12**(9):e10201. doi: 10.7759/cureus.10201.

Bhagavan NV and Ha C-E (2022) *Essentials of Medical Biochemistry*, 3rd edition. Academic Press.

Esrick EB *et al*. (2021) Post-transcriptional genetic silencing of BCL11A to treat sickle cell disease. *N Engl J Med* **384**:205–215.

Farashi S and Harteveld CL (2018) Molecular basis of α-thalassemia. *Blood Cells Mol Dis* **70**:43–53.

Garcia AA *et al*. (2021) Treatment strategies for glucose-6-phosphate dehydrogenase deficiency: past and future perspectives. *Trends Pharmacol Sci* **42**:829–844.

Graça NAG *et al*. (2022) TTP: from empiricism for an enigmatic disease to targeted molecular therapies. *Br J Haematol* **197**:156–170.

Hatton C *et al*. (2017) *Haematology Lecture Notes*, 10th edition. Wiley-Blackwell Publishing.

Hoffbrand AV and Chowdary (2024) *Hoffman's Essential Haematology*, 9th edition. Wiley-Blackwell Science.

Howard MR and Hamilton P (2019) *Hoffman's Haematology*. Churchill Livingstone.

Luzzatto L *et al*. (2020) Glucose-6-phosphate dehydrogenase deficiency. *Blood* **36**:1225–1240.

Mannucci PM and Tuddenham EGD (2001) The hemophilias – from royal genes to gene therapy. *N Engl J Med* **344**:1773–1779.

Mehta A *et al*. (2000) Glucose-6-phosphate dehydrogenase deficiency. *Baillieres Best Pract Res Clin Haematol* **13**:21–38.

Munkongdee T *et al*. (2020) Update in laboratory diagnosis of thalassemia. *Front Mol Biosci* **7**:74. doi: 10.3389/fmolb.2020.00074.

Palta S *et al*. (2014) Overview of the coagulation system. *Indian J Anaesth* **58**:515–523.

Potts DM and Potts WT (1999) *Queen Victoria's Gene: Haemophilia and the Royal Family*, 2nd edition. Sutton Publishing.

Prchal JT (2021) Chapter 51: Methaemoglobinemia and other dyshaemoglobinemias. In: *Williams Hematology*, 10th edition. Kaushansky K *et al*. (eds) McGraw Hill.

Radwan B and Bosgelmez L (2019) The safety of chelators for iron overload in sickle cell disease: a brief systemic review. *Acta Med* **50**:50–60.

Roper D *et al*. (2020) Laboratory diagnosis of G6PD deficiency. A British Society for Haematology guideline. *Br J Haematol* **189**(1):24–38. doi: 10.1111/bjh.16366.

Sundd P *et al*. (2019) Pathophysiology of sickle cell disease. *Annu Rev Pathol* **14**:263–292.

Taher AT *et al*. (2021) Molecular genetics of β-thalassemias. *N Engl J Med* **384**:727–743.

Tisdale JF *et al*. (2018) Current results of lentiglobin gene therapy in patients with severe sickle cell disease treated under a refined protocol in the phase 1 Hgb-206 study. *Blood* **132**(suppl. 1):1026. doi: 10.1182/blood-2018-99-113480.

Tisdale JF *et al*. (2020) Treating sickle cell anemia. *Science* **367**:1198–1199.

Yin T and Alquist CR (2020) Hemolytic anemia and thrombocytopenia: how to treat? *Clin Chem* **66**:763–767.

Useful websites

https://www.g6pd.org

http://scarc.library.oregonstate.edu/coll/pauling/blood/index.html

https://rarediseases.org

https://www.thalassemia.org

Taylor & Francis
Taylor & Francis Group
http://taylorandfrancis.com

Disorders of the cardiovascular system

14

OBJECTIVES

After studying this chapter you should be able to:

- outline the structure of the cardiovascular system
- describe the general functioning of the heart
- list major types of heart diseases and their causes
- review some methods for detecting and investigating heart disease
- discuss the role of cholesterol and apolipoproteins in atherosclerosis
- explain the use of antithrombolytic therapy
- describe major peripheral vascular diseases

14.1 INTRODUCTION

The cardiovascular system comprises the heart, with its enclosing pericardium, and blood vessels, that is, arteries, arterioles, capillaries, venules and veins, which enclose and distribute blood to the tissues. The heart (*Figure 14.1*) is a muscular pump weighing about 500 g in an adult and is the strongest muscle in the body. It can contract for over 100 years nonstop, beating about 100 000 times a day at a rate of approximately 70 beats/minute, to pump blood (*Chapter 13*) around the body. This supplies, for example, oxygen and nutrients to the tissues and removes their waste products. However, in such a complicated system, many things can go wrong. In the developed countries, heart disease is responsible for roughly 25% of annual deaths. Much of this is due to lifestyle, for example tobacco smoking, poor diets and obesity. In developing countries, rheumatic heart disease is more of a problem, but this has decreased in the developed countries.

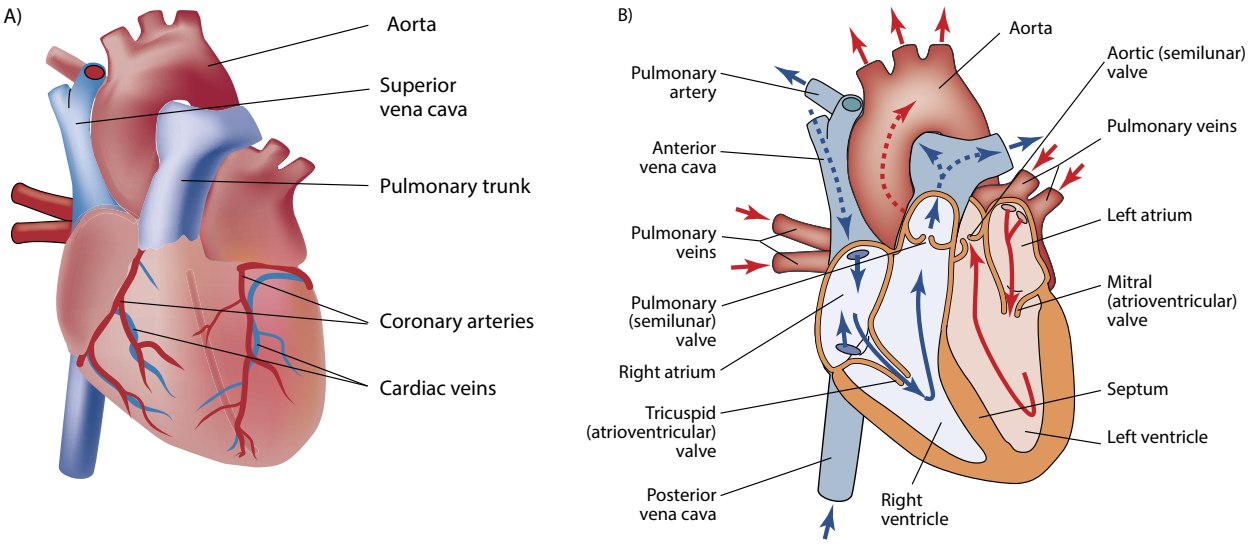


Figure 14.1 The structure of the heart. (A) Overall view of the heart and associated major blood vessels. (B) A cutaway view from the same perspective showing the heart valves and the direction of blood flow. ⏎

DOI: 10.1201/9781003296836-14

14.2 HEART AND CIRCULATORY SYSTEM

The heart (*Figure 14.1*) is a hollow, four-chambered, muscular organ situated in approximately the centre of the chest. A smooth **endocardium** lines the inside of the chambers. The wall of the heart or **myocardium** is rich in cardiac muscle tissue arranged in three layers. The inner layer is circular and thicker in the wall of the left ventricle than in the right. The outer layers of muscle spiral around the ventricles and extend to the fibrous attachments of the four valve rings. On contraction they tend to pull the chamber of the ventricle towards the valve rings. The exterior of the heart is a tough, fibrous layer that is partly covered by fat. The heart is enclosed by a double membrane system called the **pericardium**, a thin, flexible, two-layered bag that surrounds the heart. It contains just sufficient lubricating fluid between the layers so that they slide easily over each other as the heart beats. It keeps the heart in position and prevents it overfilling with blood.

Internally, the heart is centrally divided by a septum that prevents oxygenated and deoxygenated bloods from mixing. Each side is subdivided into an upper chamber or **atrium**, which collects blood and passes it to a lower chamber or **ventricle** that ejects blood. Each ventricle has a one-way inlet valve and a one-way outlet valve (*Figure 14.1*). The tricuspid valve opens from the right atrium to the right ventricle and the pulmonary valve opens from this ventricle into the pulmonary arteries. The mitral valve separates the left atrium and ventricle, and the aortic valve opens from the ventricle into the aorta. All the valves ensure that blood flows only in one direction.

Pulmonary and systemic circulations

Figure 14.2 illustrates the flow of blood through the pulmonary and systemic circulations. Oxygen-depleted blood from the body flows through the two largest veins, the venae cavae, into the right atrium. When this chamber is full, it propels the blood into the right ventricle and when full this pumps blood through the pulmonary valve into the pulmonary artery, which takes it to the lungs. As the blood flows through the network of capillaries that surround the lung alveoli it absorbs oxygen and gives up its carbon dioxide (*Chapter 9*). The resulting oxygen-enriched blood flows through the pulmonary veins into the left atrium of the heart. The complete circuit between the right side of the heart, the lungs and the left atrium is called the **pulmonary circulation**. When the left atrium is full, it sends the oxygen-rich blood into the left ventricle, which, in turn, propels the blood into the aorta, which is the largest artery in the body. From there it flows to all the tissues of the body, with the exception of the lungs, in the **systemic circulation**. The major arteries to the head, neck and arms branch off from the aorta (*Section 14.3*) and there are also smaller vessels (the coronary arteries) that supply the heart itself with blood. The aorta arches over the heart and then goes downwards carrying blood to the abdomen and legs. In the capillaries of the tissues, oxygen is exchanged for carbon dioxide, and nutrients and other materials (hormones, for example) are supplied to the tissues. The blood then returns to the heart through the veins (*Figure 14.2*).

Control of heartbeat

To work efficiently, the four chambers of the heart must beat in a coordinated way, with the atria and then the ventricles contracting simultaneously (*Figure 14.3*). A given chamber of the heart contracts when an electrical impulse moves across it. Each signal originates in a small bundle of specialised cells in the right atrium called the **sinoatrial** or **SA node** (*Figure 14.4*). This is the natural **pacemaker** that generates impulses at a given rate and ensures the heart beats regularly. Although it produces a natural rate, this can be modified by emotional and physical reactions and by hormones, especially adrenaline and noradrenaline (*Chapter 7*), which speed up the heart rate, enabling it to respond to varying demands. The electrical impulses generated by the SA node travel throughout the right and left atria causing the heart muscle to contract and arrive at the **atrioventricular** or **AV node** situated between the atria and the ventricles (*Figure 14.4*). This node delays the transmission of the impulse to allow the atria to contract completely and the ventricles to fill with as much blood as possible. The phase of relaxation of the ventricles is called **diastole**. After passing through the AV node, impulses travel through the **bundle of His**. This is a group of modified cardiac muscle fibres called **Purkinje fibres**. These divide into two branches that serve the left and right ventricles (*Figure 14.4*). The fibres spread over the surface of the ventricles in an orderly arrangement and thus initiate **systole** or ventricular contraction. Both ventricles contract together in a wringing fashion so that blood is squeezed out from their bases, pumping blood from the heart. These

Margin Note 14.1
Heart sounds

Auscultation is the use of a stethoscope to listen for distinctive diagnostic sounds within the body. If there are any abnormalities of the valves or other heart structures, then these may cause a turbulent flow of blood that creates characteristic sounds called **murmurs**, which can be heard by auscultation. Typically, the murmur is caused by blood flowing through narrowed or leaking valves. However, not all heart diseases cause murmurs and not all murmurs indicate heart disease. For example, pregnant women usually have heart murmurs because of the increase in blood flow. In infants and young children, harmless murmurs are commonly caused by the rapid flow of blood through the smaller sized heart structures.

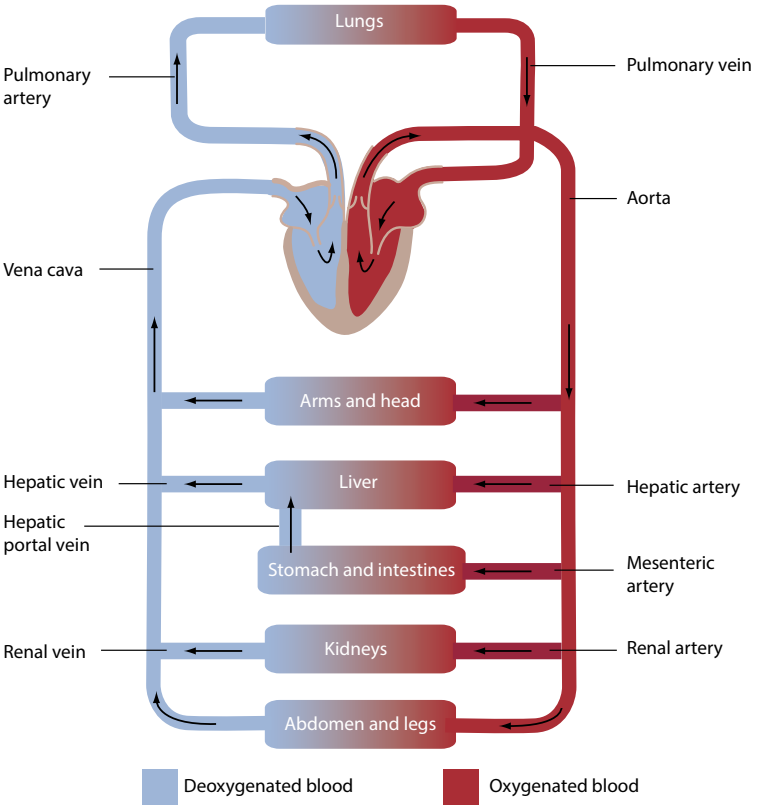


Figure 14.2 Schematic illustrating the pulmonary and systemic circulations.

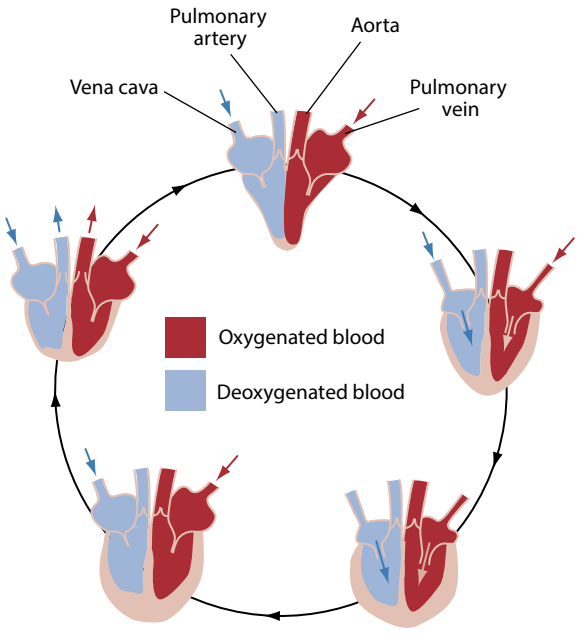


Figure 14.3 Diagrammatic representation of the contractions and relaxation of the heart chambers during one cardiac cycle. Start at the top and observe oxygenated and deoxygenated blood entering the heart, its passage from the atria to the ventricles and finally its expulsion from the ventricles as the cycle starts again.

coordinated actions keep the four chambers of the heart working in the appropriate sequence. When things go wrong, people may have to have artificial pacemakers fitted.

The part of the nervous system that regulates the heart rate automatically is, not surprisingly, the autonomic nervous system and consists of the sympathetic and the parasympathetic fibres of the vagus nerve (*Figure 14.4*). The sympathetic system increases the heart rate, while the parasympathetic system slows it down.

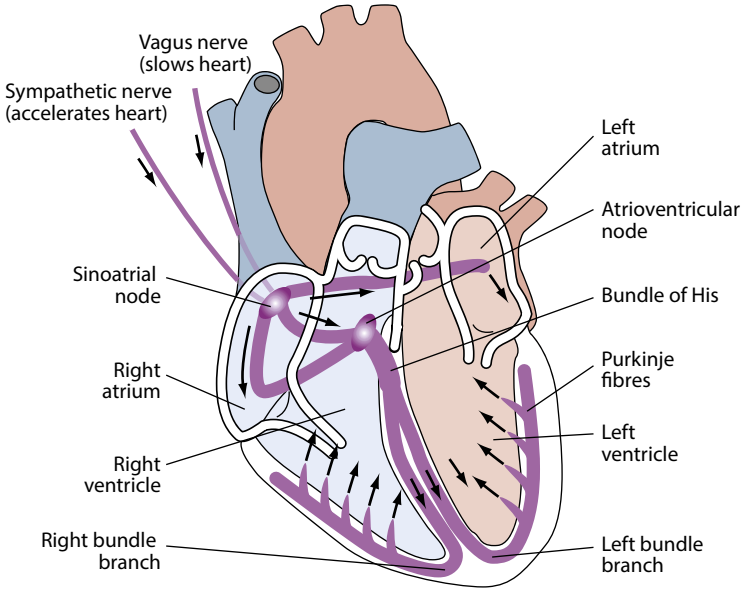


Figure 14.4 The sinoatrial (SA) and atrioventricular (AV) nodes and the control of the heart rate.

14.3 BLOOD VESSELS

The blood vessels comprise arteries, arterioles, capillaries, venules and veins (*Figure 14.5*). Arteries are strong and flexible and carry blood away from the heart. They are subjected to the highest pressure (*Section 14.4*), and their resilience helps to maintain blood pressure while the heart is between beats. The smaller arteries and arterioles have muscular walls and their diameters can be adjusted to regulate blood flow to a particular region of the body. Capillaries are thin-walled vessels that allow oxygen, nutrients and other materials, for example hormones, to diffuse from the blood to the tissues and waste products to pass from the tissues into the blood. The capillaries are links between the arteries and arterioles and the venules and veins. Veins are, in general, larger in diameter than arteries but have much thinner walls. They transport blood back to the heart but at much lower pressures and speeds than are found in arteries.

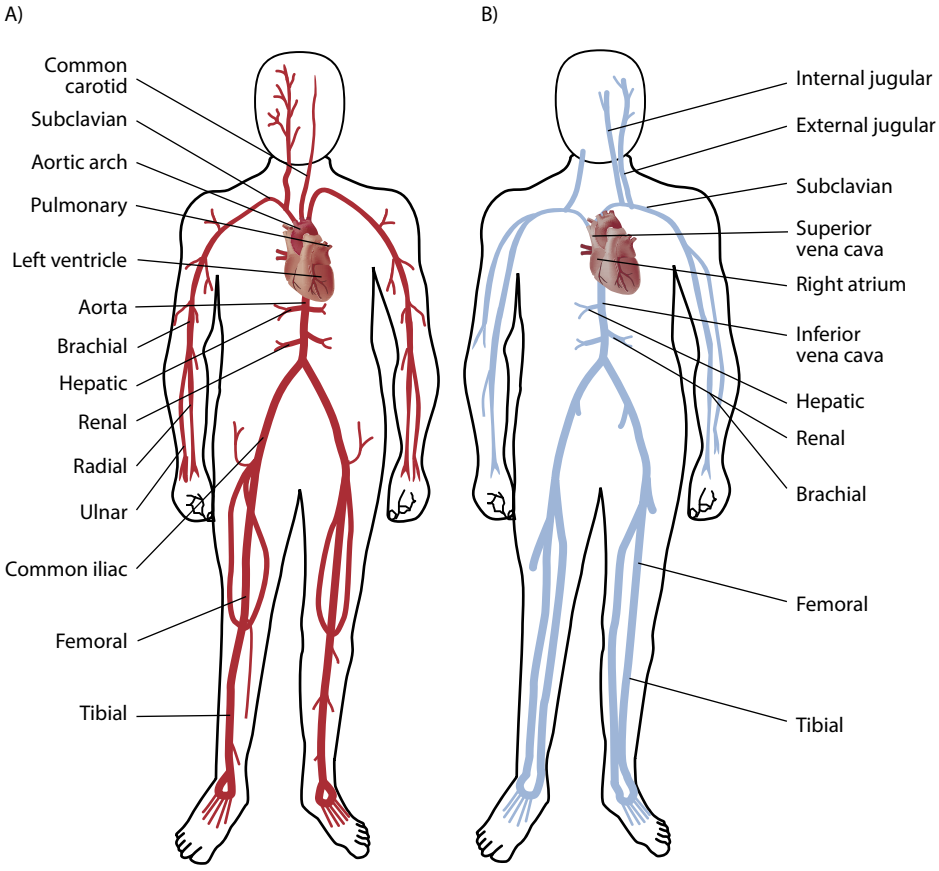


Figure 14.5 Schematic showing the main (A) arteries and (B) veins of the body. ⏎

Arteries and veins

Arteries are elastic tubes with circular cross-sections (*Figure 14.6*). They consist of three layers called tunics. The tunica intima is the layer of endothelial cells that provides a smooth, low friction innermost surface. This layer is in direct contact with the blood, and damage to it leads to serious clinical problems. The tunica media is a relatively thick middle layer composed of smooth muscle and elastic fibres, which are arranged both lengthways and circumferentially. The amount of elastic material varies with the vessel. Closer to the heart, the aorta and large branches contain relatively larger amounts, allowing them to accommodate the stroke volume. In contrast, the pulmonary artery has rather thin walls. The smooth muscle is under autonomic nervous control and regulates the diameter of the vessel. The fibrous components limit the amount of stretch under a surge of pressure. The tunica adventitia is the outer fibrous layer composed of tough collagenous material.

Arteries branch repeatedly in order to supply all parts of the body. The diameter of the aorta is approximately 25 mm and that of a medium-sized artery is about 4 mm. Junctions between the branches are called anastomoses and these may also occur between arteries and arterioles, the latter having diameters less than 0.1 mm.

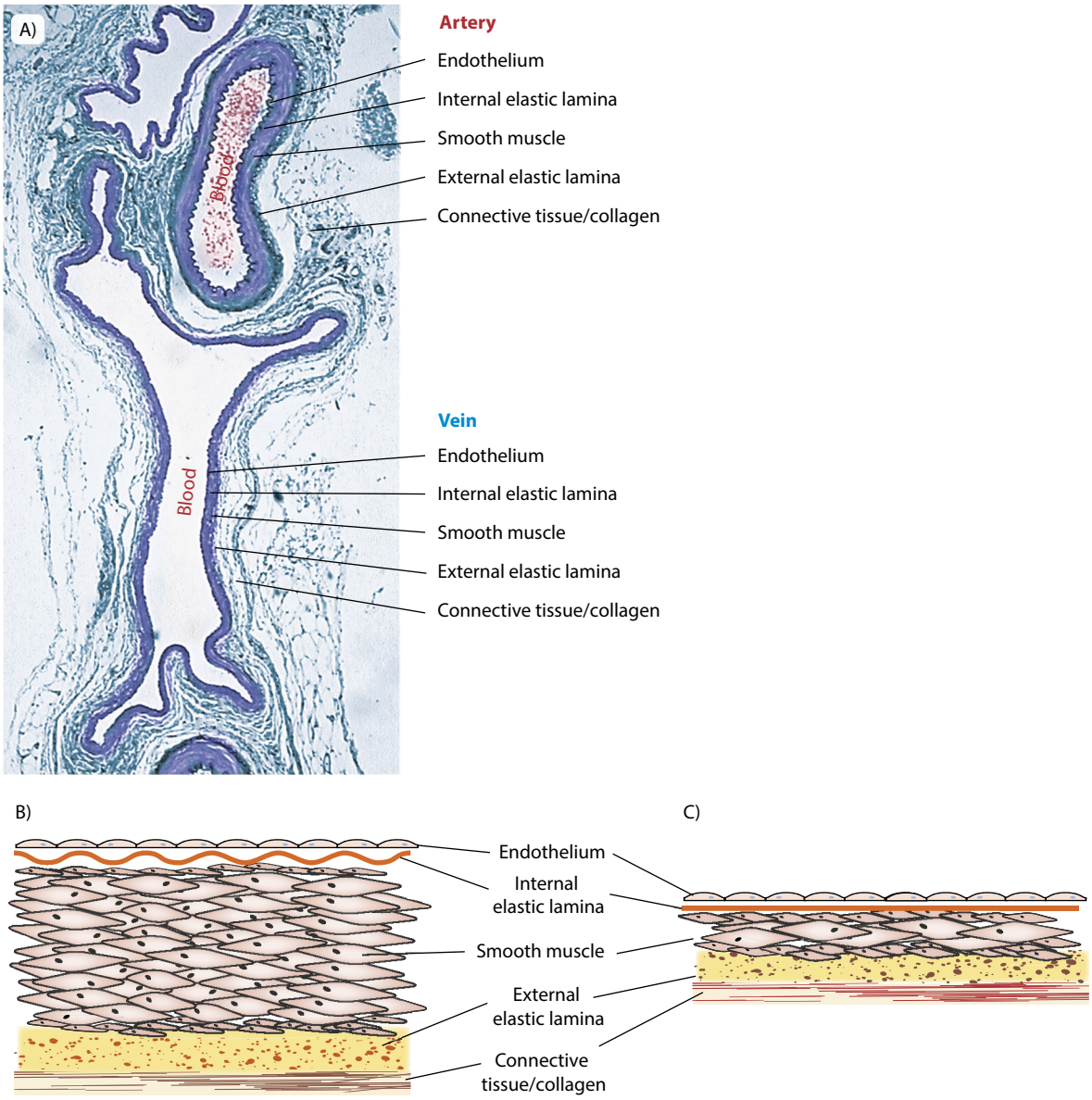


Figure 14.6 (A) Photomicrograph showing the cross-sections of an artery and a vein. Schematics illustrating the layers in the walls of (B) an artery and (C) a vein.

The pulmonary arteries carry blood from the right ventricle to the lungs. They are shorter, have thinner walls and contain blood at a lower pressure than systemic arteries. Arteries also have a storage function, since the output of the ventricles is discontinuous as the heart beats. For example, the aorta stores up blood during systole and its elastic recoil propels it on during diastole. This means that the discontinuous blood flow from the heart is converted to a more continuous one through the peripheral circulation, conserving energy and reducing the pressure that the smaller arteries have to withstand.

Veins, like arteries, have walls composed of three layers, but they contain considerably less muscle and elastic tissues and their walls are much thinner (*Figure 14.6*). The internal lining is the same endothelium as in the arteries. The veins form a blood-collecting system that returns blood from the periphery of the body to the heart. In general, veins have the same pattern of distribution as the arteries and frequently run alongside them (*Figure 14.5*). The blood collected from the systemic circuit is eventually delivered to the right atrium of the heart. However, the venous drainage from the stomach, spleen and intestine is carried by the hepatic portal vein to the liver (*Chapter 11*). Blood from the liver is then returned to the heart in the hepatic vein.

Veins are much more easily distended than arteries and are also more easily collapsed because they are less able to withstand high pressures. At intervals, especially along the lengths of long veins, the endothelial lining forms cup-shaped valves, rather similar to the semilunar valves in the heart. These allow blood to flow in only one direction, which

is back towards the heart, and also prevent its backflow under gravity. In addition, during muscular activity, especially in the legs, the compression and relaxation of the muscular tissue surrounding the veins forms a 'muscle pump' that expedites the transport of blood back to the heart. In general, the blood flow through veins is less discontinuous than it is in arteries.

Capillaries

Capillaries have internal diameters of about 7 μm, the same as that of erythrocytes (see *Figure 13.3*). Their walls are composed of a single layer of endothelial cells on a basement membrane (*Figure 14.7*) and so they lack muscle and elastic tissues. The arterioles supply beds of capillaries found deep in the tissues and these eventually deliver their blood to the venules and hence to the veins for return to the heart. The capillary network is the site where gases are exchanged, nutrients and other biomolecules are delivered, and waste products are removed. The flow of blood through capillaries is relatively slow compared with that in arteries and veins and this allows an adequate exchange to take place.

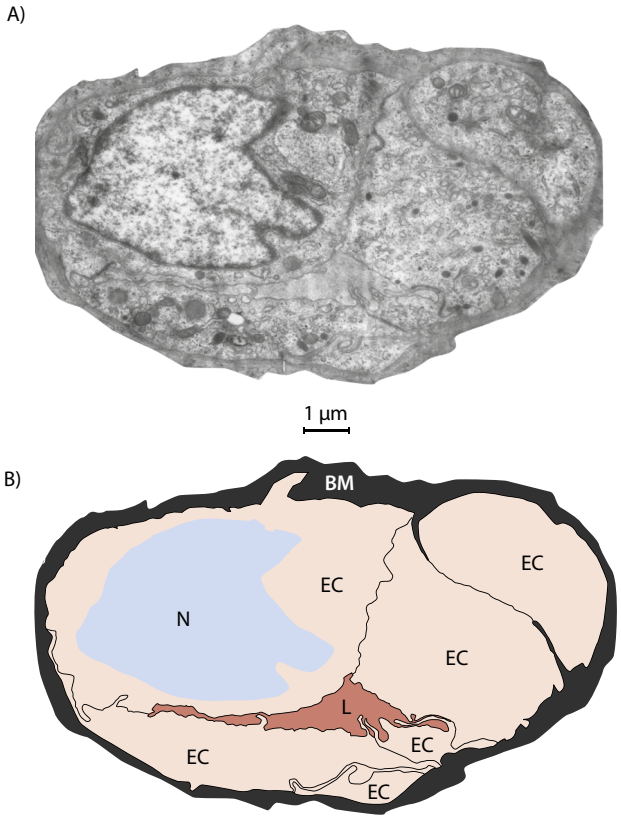


Figure 14.7 (A) Electron micrograph of a capillary and (B) its tracing. BM, basement membrane; EC, endothelial cell; L, lumen; N, nucleus of endothelial cell. Courtesy of Dr P. Kumar, School of Biology, Chemistry and Health Science, Manchester Metropolitan University, UK. ⏎

14.4 BLOOD PRESSURE

Blood pressure is the hydrostatic force that the blood exerts against the wall of a blood vessel and that propels blood around the body. It is determined partly by cardiac output and also by the peripheral resistance that occurs during diastole, when arterial walls 'snap back', but the heart contracts again before enough blood has flowed into the capillary bed to relieve the pressure. Blood pressure is highest in the arteries at systole: the systolic pressure. When the heart contracts, blood enters the arteries faster than it can leave through the capillaries, causing the vessels to stretch. This bulging of the arteries is the **pulse** and can be felt at several sites in the body. As a result, there is still a substantial blood pressure even during diastole, the diastolic pressure, and so blood flows into the arterioles more or less continuously.

Typically, at rest, the systolic and diastolic pressures are approximately 120 and 70 mmHg (16 and 9.3 kPa), respectively. Blood pressure can be measured with an automatic electronic sphygmomanometer or using a manual one combined with a stethoscope (*Figure 14.8* and *Box 14.1*). Blood pressure is not constant because, as demands on

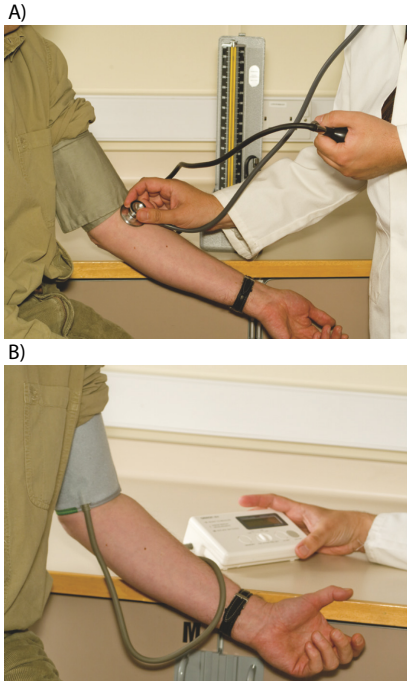


Figure 14.8 Measuring blood pressure using (A) a manual sphygmomanometer combined with a stethoscope and (B) an electronic sphygmomanometer. ⏎

BOX 14.1 MEASURING BLOOD PRESSURE

Traditionally, blood pressure is measured using a manual sphygmomanometer (from the Greek *sphygmos*, meaning pulse) and a stethoscope, as illustrated in *Figure 14.9*. There are also electronic instruments that can be used (see *Figure 14.8*). When using a manual sphygmomanometer, an inflatable cuff connected to a pressure gauge, in this case a mercury manometer, is placed around the upper arm over the brachial artery and inflated. A stethoscope is placed on the skin above the brachial artery so that blood sounds may be heard. When the pressure in the cuff is greater than the systolic pressure, the arteries are compressed and no brachial pulse can be heard. The cuff pressure is then slowly reduced and when it falls to that of the systolic pressure, clear sounds (**Korotkoff sounds**) can be heard due to the turbulence generated as the blood flows through the partially occluded artery. As the pressure further reduces, the sounds become suddenly muffled and then disappear completely. The pressure at this point is generally taken as the diastolic pressure.

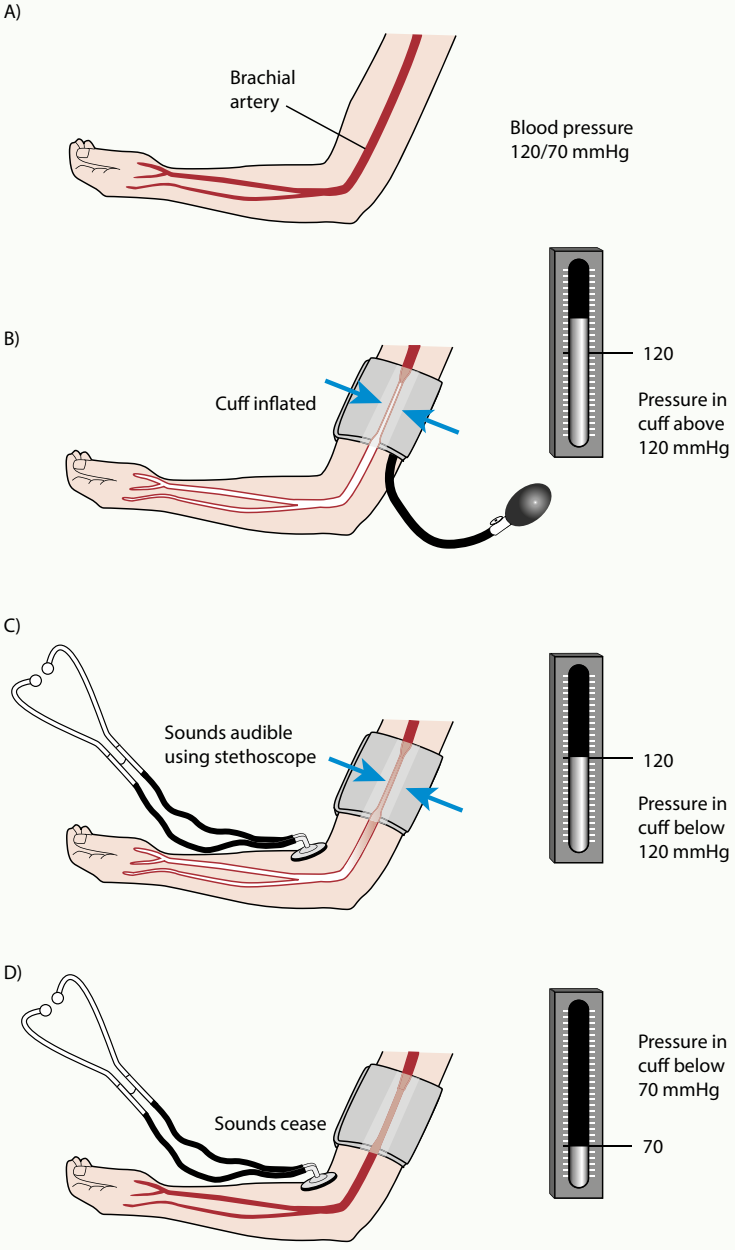


Figure 14.9 Panels (A) to (D) schematically show the use of a manual sphygmomanometer and stethoscope to measure blood pressure.

the circulatory system change, regulatory mechanisms ensure that an adequate blood flow is maintained. Thus, during heavy exercise, the arterioles in the working muscles dilate to allow increased delivery of oxygen to the muscles. This of course decreases the peripheral resistance, which, by itself, would cause a fall in blood pressure. However, the cardiac output is increased to counteract this effect.

14.5 INVESTIGATING CARDIAC FUNCTION IN HEALTH AND DISEASE

Clinicians need to be able to diagnose cardiovascular problems and to monitor their treatment(s). Obviously, noninvasive methods are to be preferred as far as possible for the comfort of the patient.

Electrocardiogram

The spread of depolarisation and repolarisation through the muscle mass of the heart is accompanied by measurable electrical potentials, which may be recorded through electrodes placed on the skin. The record of these potentials is called an **electrocardiogram** (ECG) and is the aggregate or resultant electrical activity associated with the action potentials of millions of individual cells, each of which has an amplitude and direction. Thus, recording a patient's ECG is a way of investigating the electrical activity of the heart. The test is quick, simple and painless and provides information on the heart rhythm and underlying cardiac morphology. When the pattern is abnormal it is invaluable in diagnosing a variety of heart complaints.

An ECG is recorded by placing electrodes, usually disposable, self-adhesive ones, on the chest and limbs. In practice, the ECG is recorded simultaneously from six electrodes connected to the limbs and six to the chest that measure the direction and flow of electric currents during each heartbeat. The results are recorded, with each trace representing a particular 'view' of the heart's electrical activity; these views are referred to as leads. The spikes and troughs on the graph correspond to specific events in the cycle of a heartbeat and these are lettered alphabetically. Modern ECG instruments have built-in computer chips that analyse the recordings and produce a printout of the analysis (*Figure 14.10*).

In a normal ECG waveform, the first deflection is caused by atrial depolarisation and is a low-amplitude slow deflection called a P wave. The following QRS complex results from ventricular depolarisation and, as can be seen in *Figure 14.10*, it is sharper and of a larger amplitude than the P wave. The T wave is another slow, low-amplitude wave resulting from ventricular repolarisation. Atrial repolarisation is not usually seen because it is low voltage and is hidden by the QRS complex. The PR interval is taken from the start of the P wave to the start of the QRS complex and represents the time taken for activation to pass from the SA node through the atrium to the AV node. The QRS complex is a measure of the time associated with impulses passing through the His–Purkinje system and the subsequent contractions of the ventricles. The QT interval begins at the start of the QRS complex and finishes at the end of the T wave and is the time taken to depolarise and repolarise the ventricular myocardium. The ST segment is the interval between the end of the QRS complex and the start of the T wave. At this point all the cells of the normal heart are depolarised.

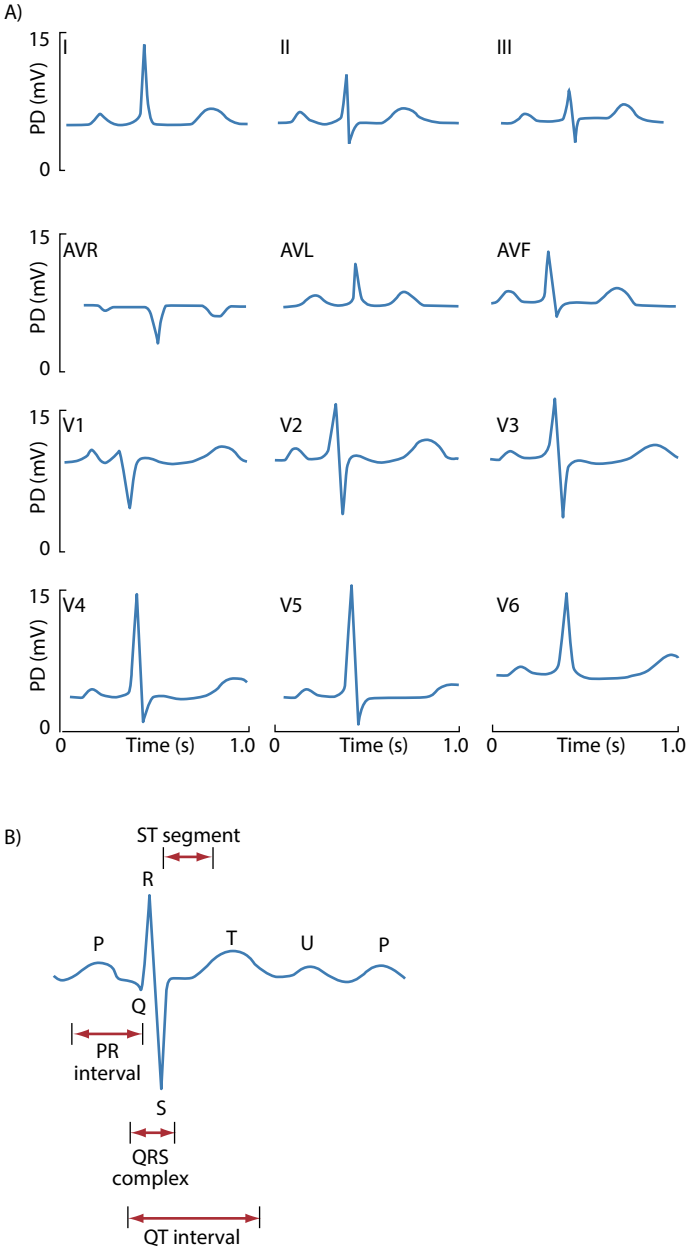


Figure 14.10 (A) Electrocardiogram (ECG) arising from the 12 leads. (B) Names of the parts of the wave observable in the graphs. See the main text for an explanation of what causes the ECG pattern. AVF, AVL and AVR, augmented limb leads, for example the AVR lead is positioned on the right arm; PD, potential difference; V1–V6, chest leads.

An ECG examination can help a cardiologist identify a number of heart problems including abnormal rhythms, inadequate blood supply to the heart and excessive thickening of the heart muscle (hypertrophy). For example, a heightened P wave indicates an enlarged atrium, a deeper than normal Q wave may indicate a myocardial infarction (*Section 14.14*) and a heightened R wave usually indicates a thickening of the ventricular wall. If the ST segment is raised above the horizontal, it may indicate acute myocardial infarction or, if below it, may imply excessive blood K^+ concentrations (*Chapter 8*) or cardiac ischaemia (*Section 14.9*). Any deviation from the normal rate or sequence of the ECG is referred to as a cardiac arrhythmia. If the SA node is damaged, the heart rate may slow to 40–50 beats/minute and if damage occurs to both the SA and AV nodes it may decrease to 20–40 beats/minute and the patient will require a pacemaker. In addition to the basic technique, cardiac responses may be recorded during exercise as the patient walks or runs on a treadmill, and this can give further clinical information. In other situations, the ECG may be recorded over 24 hours using a point-of-care device (*Chapter 1*) to monitor intermittent arrhythmias (*Section 14.7*).

Imaging techniques

Two major methods of imaging the heart are:

1. echocardiography
2. computed tomography (CT) scanning

Echocardiography uses high-frequency ultrasound to map the heart and study its various functions. It is one of the most widely used diagnostic techniques and can indicate the extent of cardiac damage. It is painless and does not use X-rays. Ultrasound waves at wavelengths of 1 mm or less (corresponding to frequencies of around 2 MHz) are generated from a probe in short bursts of a few microseconds. When the probe is pressed against the body, the emitted pulses encounter the interfaces between the various body tissues. In crossing each interface, some sound energy is reflected and is detected by a transducer in the probe and recorded as an echo. The time delays for the echoes to return are analysed by a computer to produce a moving video image of the heart and blood vessels (*Figure 14.11*). This allows a cardiologist to assess if the heart valves are functioning properly, for example whether they leak when closed, and if blood is flowing normally. Abnormal connections between blood vessels of heart chambers are also revealed, as well as contractibility.

Computed tomography is a specialised type of X-ray imaging. It gives higher resolution images than echocardiography and, indeed, more detailed information than conventional X-ray pictures (*Figure 14.12*), but is more expensive. The technique uses an X-ray machine that rotates around the patient taking sequential pictures from a number of different orientations. The images are analysed by a computer to form three-dimensional images of the area under investigation, which can be soft tissues and bones, organs, or complete cross-sections of the body.

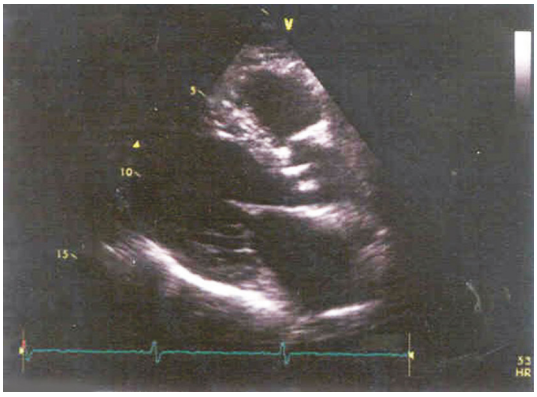


Figure 14.11 Echocardiography uses ultrasound waves that are reflected from the interfaces between the various tissues and analysed to produce the type of image shown.

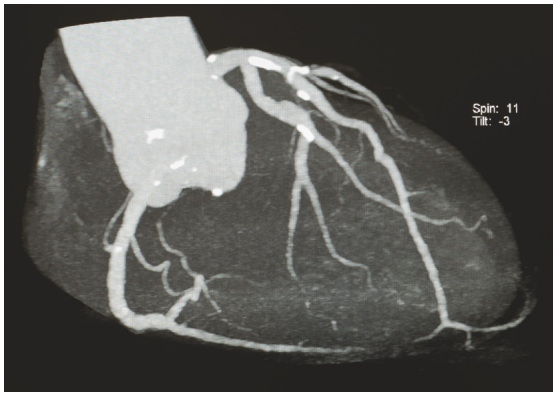


Figure 14.12 A computed tomography (CT) image of the heart. Courtesy of Shutterstock.com/Monet_3k.

Cardiac catheterisation and angiography

Cardiac catheterisation may be used to investigate heart function and abnormalities. However, it is an invasive technique and involves a certain amount of discomfort for the patient. Under local anaesthetic, a thin plastic tube is inserted into a vein or artery in an

arm or leg (groin) and advanced into the major vessels and heart chambers. The catheter can be used to measure pressure, view the inside of the heart or take blood samples to measure oxygen and lactate concentrations.

Angiography is the introduction of dyes that are X-ray opaque into the catheter, allowing X-ray investigations of possible anatomical abnormalities or abnormal blood flow. The coronary arteries can also be investigated by catheterisation to check for coronary artery disease. Chest X-rays taken from the back also enable the size of the heart in relation to the lungs and the major blood vessels to be assessed. It may also reveal areas of calcium deposition, which are a sign of tissue damage and death.

Exercise stress test

The exercise stress test measures oxygen uptake, carbon dioxide production, heart rate and lung ventilation during progressively more strenuous treadmill or cycle ergometer exercises. It can detect lung and heart diseases in their early stages and also be used to assess physical fitness. Patients with lung disease stop exercising before achieving their maximal predicted heart rate. Also, their levels of ventilation are disproportionately high for a given oxygen uptake. Hence, it is possible to differentiate breathlessness due to lung disease from that due to heart disease. For patients who cannot physically exercise, stress test measurements can be obtained using the dobutamine stress test. Dobutamine (*Figure 14.13*) is an inotropic drug that increases the heart rate; hence, giving increasingly larger doses makes the heart work harder.

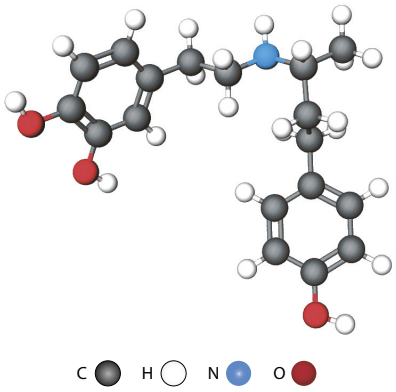


Figure 14.13 Computer drawn model of dobutamine.

14.6 ENDOCARDITIS

Endocarditis is an inflammation of the endocardium, the interior lining of the heart and its valves. It most often results from bacterial infection, which may originate from bacteria in the blood or as a result of heart surgery. The bacteria in the blood may enter from a skin wound or even from small injuries occurring when chewing food or brushing the teeth. Injecting drug users and patients with prolonged catheter use are also at risk. Patients with abnormal or damaged heart valves or fitted with artificial ones are more susceptible to endocarditis. If they are about to undergo medical or dental procedures, they must inform their surgeon or dentist and be given antibiotics prior to invasive treatments.

Bacteria and blood clots called **vegetation** can accumulate on the valves and may break loose and block vessels elsewhere in the body, causing strokes, heart attacks or pulmonary embolisms or infecting the area where they lodge.

Infective endocarditis

Acute infective endocarditis has a rapid onset and can be life threatening, unlike *subacute* infective endocarditis, which develops slowly over weeks and months. In the acute form, the symptoms are usually the sudden onset of a high fever, a fast heart rate and tiredness. There can be extensive valve damage as well as the blood clot damage elsewhere in the body. Individuals may go into shock (*Section 14.17*) and be subject to renal failure. Prompt diagnosis and hospitalisation are vital. The subacute form is associated with mild fever, tiredness, weight loss, sweating and a low erythrocyte count (*Chapter 11*). However, because the symptoms of the subacute form are more vague, damage may occur before the condition is recognised; it is just as life threatening as the acute form.

If endocarditis occurs and is identified, the treatment usually consists of at least two weeks of high-dose, intravenous antibiotics. However, heart surgery may also be necessary to repair or replace damaged valves (see *Figures 14.15* and *14.16*).

14.7 ABNORMAL HEART RHYTHMS

The contractions of heart muscle fibres and therefore the heartbeat is controlled by electrical discharges that flow through the heart along distinct pathways at controlled speeds (*Section 14.2* and *Figure 14.4*). The normal heart rate is 60–100 beats/minute, but much lower rates may be encountered and are quite normal in young adults who are physically fit. The rate, however, responds to exercise or inactivity and to pain and anger.

An inappropriately fast heartbeat is called **tachycardia** and an abnormally low rate **bradycardia**. Abnormal rhythms are frequently encountered and can be regular or irregular. If disturbances occur with the flow of the electrical discharge then **arrhythmias** in these contractions can occur. These range from the harmless to the life threatening. For example, minor arrhythmias can be caused by excessive alcohol consumption, smoking, stress or exercise. Thyroid hormones also affect the heart rate and an over- or underactive thyroid gland (*Chapter 7*) may affect the rate and rhythm of the heart. Some of the drugs used to treat lung disease or high blood pressure can have similar effects.

The most common causes of arrhythmia are heart disease, especially coronary heart disease, heart failure or abnormal valve function. In many cases, patients are aware of an abnormal heartbeat and this is referred to as **palpitations**. This awareness may be disturbing but there are many possible causes of arrhythmias and often they are not the result of an underlying disease. However, when they are, it is the nature and severity of the disease that is more worrying than the arrhythmia itself. A clinician would want to determine if the arrhythmias are fast or slow, regular or irregular; whether they make the person feel dizzy, light-headed or even lose consciousness; whether they occur at rest or during exercise; and if they are accompanied by a shortness of breath or chest pain.

An ECG is, of course, helpful in diagnosis, but an arrhythmia may only occur over a short period. Consequently, a portable monitor may be used to record heart rhythms over 24 hours. The prognosis and treatment will depend on whether the arrhythmias start in the pacemaker, the atria or the ventricles. Most arrhythmias are harmless and do not interfere with the heart's pumping action. However, antiarrhythmic drugs are available if the patient is anxious or if the arrhythmias cause intolerable symptoms or pose a risk. Sometimes it is necessary to fit the patient with an artificial pacemaker that is programmed to replace the heart's own pacemaker. These are usually implanted surgically under the skin of the chest and are wired to the heart. Most commonly they are used to correct abnormally slow heart rates. Sometimes an externally applied electric shock to the heart, called cardioversion, defibrillation or electroversion, can correct an abnormal arrhythmia. Arrhythmias caused by coronary artery disease may be controlled by drugs, pacemakers or surgery. After a coronary infarction, some people have life-threatening episodes of ventricular tachycardia that may be triggered in an injured area of heart muscle that may have to be removed during open heart surgery.

Atrial fibrillation and flutter

Atrial fibrillation and flutter are rapid electrical discharges that make the atria contract rapidly, but each contraction may not conduct to the ventricles. This causes the ventricles to contract irregularly and less efficiently, producing a condition that may be sporadic or persistent. During the fibrillation or flutter, the atrial walls simply quiver and blood is not pumped into the ventricles properly. The consequence is that inadequate amounts of blood are pumped from the heart, blood pressure falls and heart failure may occur. The diminished pumping ability may make the patient feel weak, faint and short of breath. Sometimes, especially with elderly patients, there is chest pain and heart failure. If the atria do not contract completely, blood may stagnate in the atria and clot. If pieces of clot or emboli break off, they may move and block an artery elsewhere in the body. If this is in the brain it may cause an embolism or stroke (*Section 14.16*).

The treatment for atrial fibrillation is to correct the disorder that causes the abnormal rhythm, restoring it to normal, and to slow the rate at which the ventricles contract so as to improve the pumping efficiency. The latter can often be achieved with digoxin, but a β blocker, such as atenolol, bisoprolol, metoprolol or propranolol (*Figure 14.14*), may also be required. Frequently, a normal rhythm must be restored by antiarrhythmic drugs or cardioversion, which is the use of a low-energy electric shock administered by

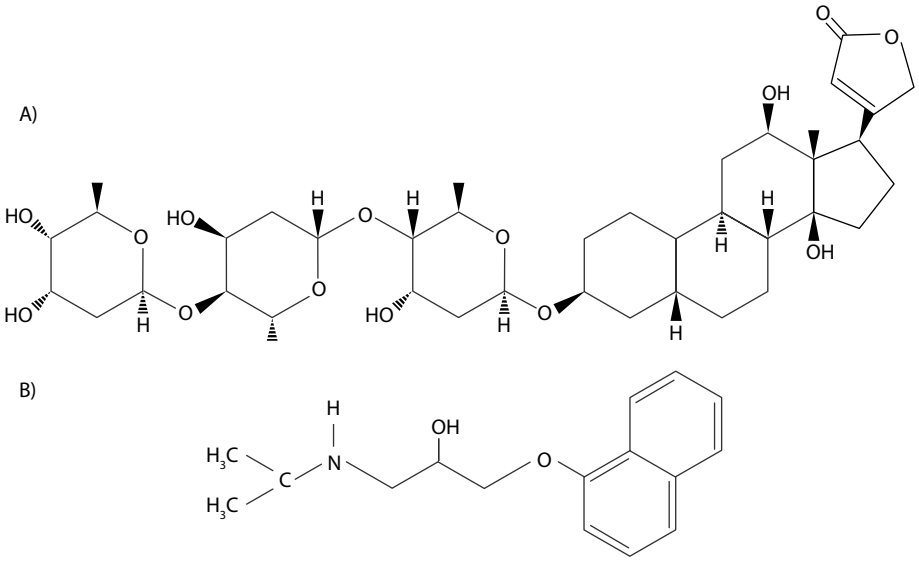


Figure 14.14 (A) Digoxin and (B) propranolol. See also *Figure 14.34*.

a defibrillator and synchronised to the QRS complex (*Figure 14.10*). In these conditions, blood can pool in the ventricles and clot; hence, anticoagulant drugs may be used when atrial fibrillation is present.

Ventricular tachycardia

Sustained ventricular tachycardia, with a ventricular rate of at least 120 beats/minute, occurs in various heart diseases that damage the ventricles. Most often it occurs over weeks or months after a heart attack (*Margin Note 14.2*). It is characterised by palpitations and will usually require emergency treatment because the blood pressure falls and heart failure may follow. Cardioversion is needed immediately.

Ventricular fibrillation

Ventricular fibrillation is a form of cardiac arrest. It is similar to atrial fibrillation, but the prognosis is serious and potentially fatal if not treated immediately. It is the product of an uncoordinated series of rapid but ineffective contractions throughout the ventricles. These, in turn, arise from multiple chaotic electrical impulses. Its most common cause is an insufficient flow of blood to the heart muscle because of coronary heart disease or a heart attack (*Margin Note 14.2* and *Section 14.14*). Given that blood is not pumped from the heart, it can lead to unconsciousness in seconds and, if untreated, the patient usually has convulsions and develops irreversible brain damage because of oxygen starvation. Thus, ventricular fibrillation is a medical emergency. Cardiopulmonary resuscitation must be started in as short a time as possible, usually within three minutes. This should be followed by cardioversion also as soon as possible. Subsequently, drugs are needed to restore and maintain the normal heart rhythm.

Heart block

Heart block is a delay in electrical conduction through the AV node. It has various degrees of seriousness; the least may not need treatment, but the most serious may require the fitting of an artificial pacemaker. First-degree heart block is relatively common in well-trained athletes, teenagers and young adults, but it may also be caused by rheumatic fever (*Box 14.2*) or by certain drugs. At the other extreme, in third-degree heart block, electrical impulses from the atria to the ventricles are completely blocked. The ventricles beat so slowly that their pumping ability is compromised. Fainting, dizziness and sudden heart failure are common.

Margin Note 14.2
Heart attack or myocardial infarction

A heart attack is a **myocardial infarction** (*Section 14.14*). It is a medical emergency during which some or all of the blood supply to the heart muscle through the coronary arteries is cut off. If this continues for more than a few minutes, the heart tissue dies, with serious clinical consequences or even the death of the individual. The term **heart attack** is sometimes used loosely for other heart conditions: strictly, it should only be applied to myocardial infarctions.

BOX 14.2 RHEUMATIC FEVER

Rheumatic fever occurs mostly in children and young adults and is caused by infection with group A *Streptococci*. It is now much less common in the developed countries than was previously the case: for example 10% of children in the 1920s compared with about 0.01% now. This is mainly due to the development of antibiotics. Rheumatic fever is thought to result from an autoimmune reaction triggered by the bacteria rather than any bacterial toxin (*Chapters 2* and *3*). The skin, joints and central nervous system and all the layers of the heart may be affected. The disease presents with fever, joint pains, malaise, loss of appetite and a characteristic fleeting polyarthritis affecting the larger joints, such as knees, elbows and ankles, which become swollen, red and tender. Effects on the heart include new or changed murmurs, cardiac enlargement or failure, and pericardial effusion.

Usually, nonspecific indicators of inflammation, such as the erythrocyte sedimentation rate and C-reactive protein (*Chapter 13*), may be elevated. Throat swabs should be cultured to investigate for group A *Streptococci* and there may be serological changes indicative of a recent streptococcal infection. If patients have fever, active arthritis or active **carditis** (*Section 14.6*), that is, a general inflammation of the heart, they should be completely bed rested. Residual streptococcal infection should be eradicated with a single intramuscular injection of benzathine penicillin or four daily oral doses of phenoxymethyl penicillin for a week. Salicylate and steroids may also be given if carditis is present. Recurrence is common. More than half of patients with acute rheumatic fever with carditis will develop conditions after 10–20 years that affect the mitral and aortic valves.

14.8 CARDIAC FAILURE

Cardiac failure is the inability of the heart to maintain an adequate cardiac output, that is, pump a volume of blood per minute sufficient to meet the demands of the body. The heart does not stop beating, as is often thought, but its diminished ability imposes severe demands. It is a serious condition, but is more common in older people. The incidence is about 1 in 100 for individuals over 65 years. Irrespective of its cause, the prognosis is

relatively poor: over 50% of patients will survive for 5 years and 35% will live for 10 years following diagnosis. New drugs, such as dapagliflozin, may improve prognosis.

There are many possible causes of heart failure and, indeed, any disease that affects the heart and circulatory system can lead to heart failure. The most common of these is coronary artery disease, which limits the flow of blood and, hence, oxygen and nutrients to the heart muscle and can lead to a heart attack. Bacterial and viral infections (*Chapter 2*) can also damage the heart muscle, as can diabetes, an overactive thyroid (*Chapter 7*) and obesity (*Chapter 10*). Obstruction of the heart valves or heart valves that leak increases its workload, and this eventually weakens the contractions. Similarly, a narrowed aortic valve means that the heart has to work harder because it has to force blood through a smaller exit, again imposing an extra metabolic burden. High blood pressure (*Section 14.17*) also means that the heart has to work too hard. Diseases that affect electrical conduction in the heart can result in an abnormal heartbeat that reduces pumping efficiency. Although the increased workload initially results in hypertrophy, that is, enlargement, of the heart muscle so that it can contract with greater force, eventually the heart malfunctions, making the heart failure worse.

Heart failure results in tiredness and weakness during physical activities because the skeletal muscles are starved of blood. The disease may be on one side of the heart or the other, but the condition usually affects the whole heart. Nevertheless, there are characteristic symptoms depending on which side is affected. Right-sided disease tends to cause a build-up of blood flowing into the right side of the heart, which leads to swelling of the feet, ankles, legs and liver. In contrast, left-sided disease increases fluid in the lungs (pulmonary oedema), causing, in turn, shortness of breath. At first this is only experienced during exertion, but it gradually increases in severity so that the breathlessness occurs even at rest. If this happens at night, the patient may wake up gasping for breath and may find it better to sleep in a sitting position. Cardiac failure gradually worsens with time if the underlying disease is not treated, although patients may continue to live for many years.

The symptoms described above are usually sufficient for an initial diagnosis of heart failure, which would be confirmed by a weak and rapid pulse, lowered blood pressure and abnormal heart sounds. However, its underlying cause must also be identified. In many cases, taking a clinical history and examining the patient will be sufficient. General diagnostic tests include chest X-ray to demonstrate an enlarged heart and fluid accumulation in the lungs, ECG, echocardiography and blood tests (for example full blood count, liver function tests [*Chapter 11*] and determination of blood concentrations of urea and electrolytes [*Chapters 11* and *13*]), and analysis of cardiac biomarkers in acute heart failure to diagnose myocardial infarction (*Section 14.14*) will then be performed. Functional tests may also be performed, including exercise testing, ECG monitoring and angiography at rest and under stress. The treatment for heart failure is focused on relieving the symptoms, retarding the progression of the disease and aiming to improve the chances of survival. This means that any factor aggravating the failure should be identified and treated. The precise cause of failure should be identified and if possible corrected. For example, if the cause is a narrowed aortic valve, treatment may require replacing it with an artificial one (*Figure 14.15*). Patients should be nursed in a comfortable, upright position.

Figure 14.15 An artificial aortic valve. Courtesy of Shutterstock.com/Mehmet Cetin.

Chronic heart failure

In chronic heart failure, the circulation at rest is adequate but there is an inadequate reserve to pursue daily activities. Its treatment depends upon the underlying disease to be dealt with. For example, heart surgery can correct narrowed or leaking heart valves (*Figure 14.16*), and bypass surgery (*Section 14.14*) can alleviate blocked coronary arteries. If the disease is caused by an infection, the condition may be improved by antibiotics without the need for surgery. Additionally, there are many things a patient can be advised to do to help the condition, including giving up smoking, eating less salt, reducing excessive weight and controlling alcohol consumption.

The best treatment for heart failure is to prevent it happening in the first place or by reversing its underlying cause as soon as possible. Nevertheless, there is still much that can be done. For example, if reducing the salt intake does not lower fluid retention, then diuretics may be prescribed. A reduction in the amount of body fluid reduces the volume of blood to be pumped and so alleviates the strain on the heart. Digoxin (*Figure 14.14*) may be given to increase the power of each heart contraction and will slow a rate that is too rapid. Vasodilatory drugs, such as angiotensin-converting enzyme (ACE) inhibitors (*Margin Note 14.3*), hydralazine and minoxidil, may be prescribed to expand blood vessels, lowering the blood pressure. Some of the older drugs dilate arteries more than veins or vice versa. However, ACE inhibitors dilate both arteries and veins and are, perhaps, the most commonly used. These improve the symptoms and prolong life. A heart

Figure 14.16 An artificial heart valve. Courtesy of Shutterstock.com/BelezaPoy.

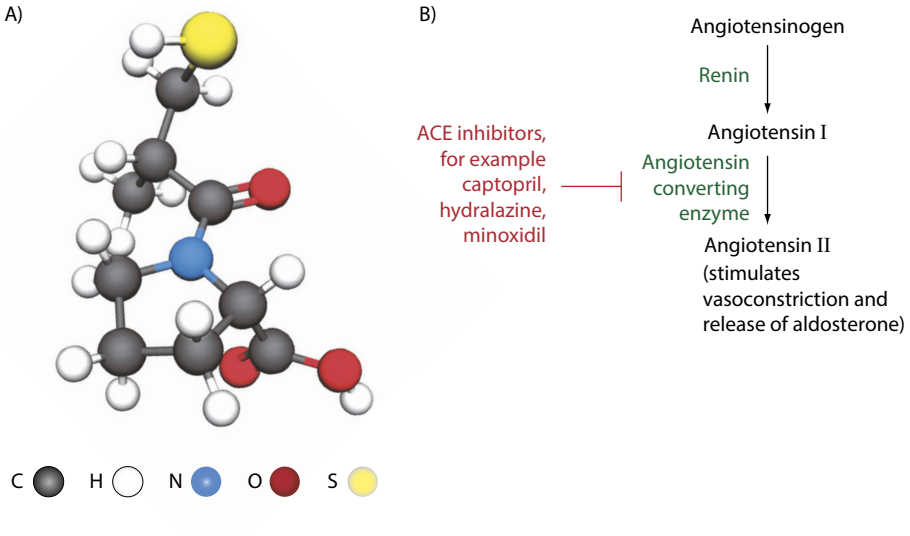


Figure 14.17 (A) Computer drawn model of the angiotensin-converting enzyme (ACE) inhibitor drug captopril. (B) An outline of the inhibition of ACE by the action of ACE inhibitor drugs.

transplant is perhaps the ultimate possibility but, sadly, there are never enough good hearts to go around (*Chapter 6*).

Acute heart failure

In acute heart failure, the haemodynamic derangement is so severe that it results in symptoms even at rest. If fluid suddenly accumulates in the lungs, that is, acute pulmonary oedema, and the patient has to gasp for breath, then emergency treatment is required. Oxygen is given by a face mask and together with intravenous diuretics may result in a rapid and dramatic improvement. Glyceryltrinitrate (*Figure 14.18*) may be given intravenously or placed under the tongue, and this leads to dilation of the veins, reducing the amount of blood flowing through the lungs. It may be necessary to insert a tube into the patient's airway to help breathing. The treatment for acute heart failure is essentially the same as for the chronic condition, as described above.

14.9 ISCHAEMIC HEART DISEASE

Myocardial ischaemia, meaning a lack of oxygen to the myocardium, results from an imbalance between the demands of the myocardium for oxygen and the amount supplied. It has several possible causes. Firstly, the blood flow through the coronary arteries may be reduced because of mechanical blockages, such as atheroma, a plaque or thrombosis (a clot); spasm (a sudden muscular contraction); an ostial stenosis (a narrowing or occlusion of the coronary ostium); arteritis (that is, an inflammation of the blood vessel); or any sort of blockage (an embolism) due to, for example, tumour cells or an air bubble. Secondly, there may be a decreased flow of oxygenated blood to the myocardium because of anaemia (*Chapter 13*), **hypotension** (low blood pressure) or carbon monoxide poisoning (*Chapter 12*). Thirdly, an increased demand for oxygen can be the result of exercise or myocardial hypertrophy that require increases in cardiac output. Note that, in the last two scenarios, the coronary arteries may be healthy. The most common cause of ischaemic heart disease is coronary atheroma, which obstructs the flow of blood through the coronary arteries.

Diagnosis and treatment of ischaemic heart disease

A diagnosis of ischaemic heart disease is based on clinical history (that is, the presentation of symptoms and family medical history), a physical examination and tests such as ECG, chest X-ray, echocardiogram, magnetic resonance imaging (MRI; see *Chapter 18*) or CT scans, and cardiac catheterisation. Sometimes a myocardial biopsy may be performed to determine the cause of ischaemic heart disease. Treatments consist of general advice to the patients, including suggestions for changing diet and lifestyle, such as not smoking, eating healthily and weight management (*Chapter 10*), blood pressure control, and regular exercise. Antiplatelet therapy should be prescribed for all patients, and dual antiplatelet therapy may be considered in selected patients. β-Blockers are often used to treat patients with ischaemic heart disease.

Margin Note 14.3
Angiotensin-converting enzyme inhibitors

Angiotensin-converting enzyme (ACE) inhibitors are used to treat chronic heart failure and high blood pressure. The drugs block the conversion of angiotensin I to angiotensin II, which is both a vasoconstrictor and the most important stimulus for the release of aldosterone (*Chapter 8*) from the adrenal cortex. Angiotensin-converting enzyme inhibitors are effective antihypertensive drugs because they lower the systemic vascular resistance and venous pressure and reduce concentrations of circulating catecholamines, thus improving myocardial performance. The first ACE inhibitor used was captopril, but commonly used ones include enalapril, lisinopril, perindopril and ramipril. Their mode of action is well understood (*Figure 14.17*). If the flow of blood through the kidneys is low, they release renin (see *Figure 8.2*), which converts angiotensinogen to angiotensin I. Subsequently, ACE converts angiotensin I into angiotensin II. Among its other effects on blood vessels, angiotensin II causes the division of heart muscle cells and fibrosis, which can make heart failure more serious in the long run. Giving ACE inhibitors reduces or eliminates these effects. Angiotensin-converting enzyme inhibitors are used routinely post myocardial infarction to maintain good cardiac action and prevent heart failure.

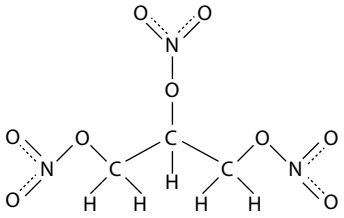


Figure 14.18 Glyceryltrinitrate.

Patients who still have angina despite lifestyle changes and medical therapy may warrant percutaneous or surgical revascularisation to reduce symptoms.

14.10 CARDIOMYOPATHIES

Cardiomyopathy is a progressive disorder that impairs the function of the ventricular muscle walls. It may be idiopathic or result from a number of clinical conditions, including:

- dilated congestive cardiomyopathy
- hypertrophic cardiomyopathy
- restrictive cardiomyopathy

Dilated congestive cardiomyopathy

Dilated congestive cardiomyopathy is not a single condition but a group of heart disorders in which the ventricles have enlarged but are still not able to pump enough blood to meet bodily needs. Heart failure may result. The most common cause of the defect in developed countries is widespread coronary artery disease, which leads to an inadequate blood supply to the heart muscle. This causes damage, and the undamaged muscle then stretches in compensation. If this is inadequate to meet body needs, then dilated congestive cardiomyopathy develops. Its symptoms are shortness of breath on exertion and a rapid onset of tiredness due to a weakened pumping action. The heart rate speeds up, so blood pressure is normal or low, but fluid is retained in the legs, abdomen and lungs. Enlargement of the heart can mean that the valves do not close properly, leading to leakage, and this improper closing may be heard as murmurs using a stethoscope (*Box 14.1*). The stretching may also increase the potential for arrhythmias (*Section 14.7*). Electrocardiography or MRI may be used to confirm the initial diagnosis.

Patients with the condition have a poor prognosis and, indeed, the outlook declines as the heart walls become thinner and contractibility is reduced: one year mortality rates are 25–30% and for five years are approximately 50%. Men tend to survive only half as long as women, and Black people half as long as the White population. In about half of the cases there is sudden death. Treating the underlying cause, for example reducing alcohol abuse, can prolong life. When there is coronary artery disease there may be angina (*Section 14.13*), which is treated with glyceryltrinitrate, β-blockers or calcium channel antagonists. There may also be pooling of blood in the swollen heart that can cause clots to form and therefore the patient is given anticoagulants (*Chapter 13* and *Box 14.3*).

BOX 14.3 ANTICOAGULANT THERAPY ⏎

Venous thromboembolism is a common problem after surgery, especially in patients who are elderly, have malignancies or have a previous history of thrombosis. There is also a high incidence in patients confined to bed after trauma or a variety of heart diseases. Its prevention and treatment include the use of anticoagulants, such as heparin, warfarin and aspirin. Heparin used for therapeutic purposes is usually produced from porcine intestines. However, there is considerable interest in heparin from other sources in case of a shortage in the global supply of porcine heparin. Heparin is a mixture of polysaccharides (but see *Figure 13.24*) that potentiates the formation of irreversible complexes between antithrombin III (a potent inhibitor of coagulation) and several coagulation factors (for example thrombin and factors VIIa, IXa, Xa, XIa and XIIa [*Chapter 13*]). Its injection has an immediate effect on blood coagulation, but its action is relatively short lived. Bleeding complications of heparin treatment occur occasionally and are treated by giving the positively charged protein protamine. Low-M_r heparins, produced by chemical degradation of standard heparin, have somewhat different properties. Their longer half-life means they can be given as a once- or twice-daily subcutaneous injection, which is preferred, or as a continuous infusion, not the six-hourly injection, as is required with regular heparin. They also seem to cause less inhibition of platelet function and thus there is a reduced risk of causing bleeding.

Oral anticoagulants work by interfering with phylloquinone or vitamin K metabolism (*Chapters 10* and *13*). Vitamin K is a cofactor required for the formation of γ carboxyglutamate residues in prothrombin (*Chapter 13*). The coumarin anticoagulant warfarin is frequently used because of its low incidence of side effects other than promoting bleeding. However, many drugs

interact with warfarin. Tricyclic antidepressants (such as amitriptyline), clofibrate, aspirin and alcohol all increase the anticoagulant effect of warfarin, while other drugs such as rifampicin and barbiturates decrease its effects. Contraindications include pregnancy, peptic ulcers, severe liver and renal diseases, and preexisting haemostatic disease.

Low doses of aspirin inhibit platelet (see *Figure 13.8*) aggregation by blocking cyclooxygenase activity irreversibly, thus altering the balance between the complex fatty acids, prostacyclins and thromboxane A_2, which is required for platelet activation (*Figure 14.19*). Long-term treatment with aspirin greatly reduces the risk of myocardial infarction (and death) in patients with angina.

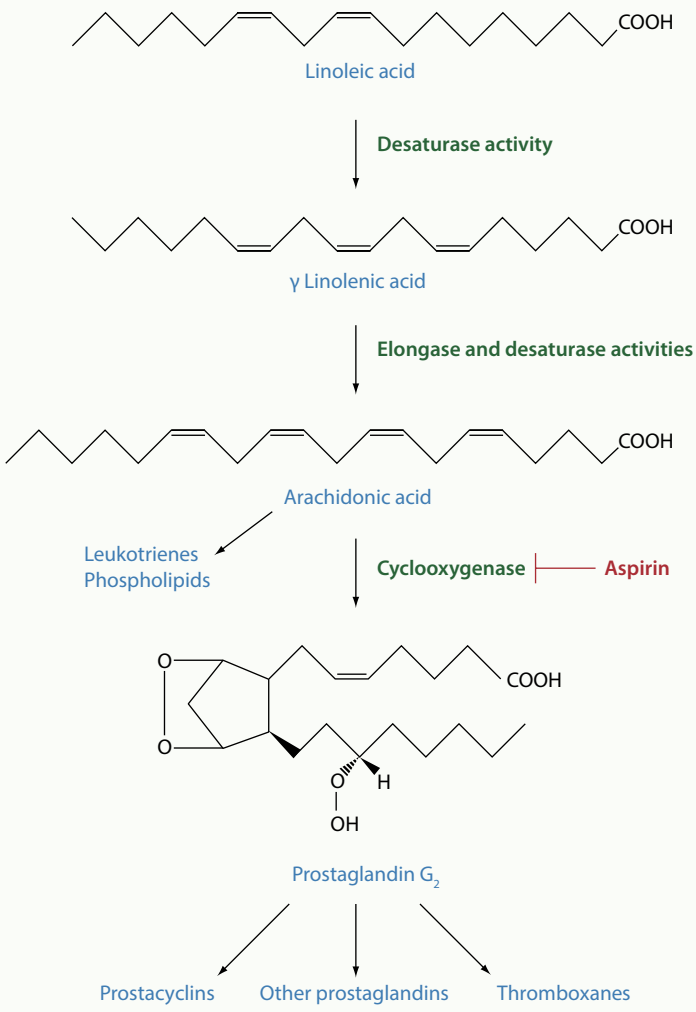


Figure 14.19 Outline of the inhibition of cyclooxygenase activity by aspirin, which disrupts the biosynthetic balance between complex fatty acids.

Myocarditis or viral cardiomyopathy

Myocarditis, an acute inflammation of the heart muscle, is caused by a viral infection, most often coxsackie virus B (see *Figure 3.1*); hence, it is sometimes called viral cardiomyopathy. It may weaken the heart muscle, producing a condition similar to dilated congestive cardiomyopathy (see above). A number of chronic hormonal disorders (*Chapter 7*), including diabetes and thyroid disease, can also produce cardiomyopathy, as can prolonged alcohol abuse (*Chapter 12*).

Hypertrophic cardiomyopathy

Hypertrophic cardiomyopathy is a group of conditions in which the ventricular walls thicken. It may occur as a birth defect, in adults with acromegaly or in people with pheochromocytoma, a tumour of the adrenal gland (*Chapter 7*). The heart becomes thicker

and stiffer than normal and more resistant to filling with blood from the lungs, leading to a back pressure on the lung veins, causing a pulmonary oedema. The patient therefore becomes chronically short of breath, with symptoms that include faintness, chest pains and palpitations brought on by irregular heartbeats; the heart sounds through a stethoscope are usually characteristic. Previously, younger patients tended to die suddenly of hypertrophic cardiomyopathy, but with better and earlier diagnosis and drug therapies this is now less of a problem.

The main treatments are administering β-blockers and calcium channel inhibitors. Surgery to remove some of the heart muscle may improve the outflow of blood, but this is only undertaken when the patient is incapacitated despite drug therapy.

Restrictive cardiomyopathy

Restrictive cardiomyopathy is the least common of the cardiomyopathies. It has many features in common with hypertrophic cardiomyopathy. The replacement of heart muscle tissue with scar tissue or its infiltration with abnormal material, such as leukocytes, amyloidosis and sarcoidosis (*Box 14.4*; see also *Box 18.2*), can all lead to the condition, although its cause is frequently unknown. The major symptoms are shortness of breath and an oedematous swelling of the tissues. About 70% of individuals with the disease die within about five years of the symptoms commencing. In most cases, therapy is unsatisfactory. The use of diuretics, which are normally given to treat heart failure and reduce the amount of fluid accumulating in the tissues, may actually reduce the amount of blood entering the heart and worsen the condition.

BOX 14.4 AMYLOIDOSIS AND SARCOIDOSIS ⏎

There are several forms of amyloidosis in which amyloid protein that is not normally present in the body, and which is of unusual conformation, accumulates in various tissues (see *Box 18.2*). One type of amyloidosis affects the heart and is associated with normal ageing, but what causes its build-up is not known with any certainty. The condition may not need treatment and, in any case, treatments are not usually very successful. A patient with heart problems may be given a heart transplant (*Chapter 6*), but the transplanted organ itself may be affected later. In contrast with amyloidosis, sarcoidoses are collections of abnormal inflammatory cells called granulomas. They commonly appear in the lymph nodes, lungs, liver, eyes and skin, but can also occur in the spleen, bones, joints, skeletal muscles, heart and nervous systems. Their cause is varied, but they seem to result from an abnormal response of the immune system (*Chapter 4*). In some instances, there are no symptoms, while in others there can be fever, weight loss and aching joints. Often the condition clears up spontaneously and the granulomas may eventually disappear or become scar tissue. In other cases, there may be permanent damage, such as lung scarring. Most people do not need treatment, but corticosteroids (*Chapter 7*) may be used to suppress severe symptoms, such as shortness of breath or severe skin lesions.

14.11 HEART VALVE DISORDERS

Major problems encountered with heart valves include an improper closing of the valves leading to leakage (**regurgitation**) or failure to open fully (**stenosis**). These conditions interfere with the heart's capacity to pump blood. If the mitral valve leaks, then regurgitation occurs each time the ventricle contracts. As blood is pumped into the aorta, some leaks back into the left atrium, increasing the volume and the pressure in that compartment. This, in turn, increases the blood pressure in the vessels leading from the lungs to the heart, resulting in a pulmonary oedema. Rheumatic fever was once the most common cause of mitral valve regurgitation (*Box 14.2*), but heart attacks that damage the structures supporting the valve are now the most common cause. However, in countries where there is poor preventative medicine, rheumatic fever is still common. Repair or replacement of the valve is required if the regurgitation is severe (*Figure 14.16*).

The aortic valve may also become leaky, and the most common causes were rheumatic fever and syphilis, but these are now more rare because of the use of antibiotics. In contrast, aortic valve stenosis is mainly a disease of the middle aged and elderly (over 60 years) and is the result of the valve becoming calcified. The left ventricle wall thickens as the heart strains to pump sufficient blood through the narrowed opening and the

enlarged heart muscle requires extra oxygen and nutrients from the coronary arteries. Eventually, the output of blood from the heart becomes insufficient to meet physiological requirements, and the resulting heart failure causes shortness of breath and fatigue. The treatment is to replace the aortic valve, preferably before the left ventricle becomes irreparably damaged.

Clinical problems may also occur with the tricuspid valve. However, regurgitation usually requires little treatment. Stenosis of the tricuspid valve is rare, again because the damage is mainly associated with rheumatic fever.

14.12 HEART TUMOURS

Heart tumours (*Chapter 17*) may be asymptomatic or precipitate life-threatening crises, such as sudden heart failure, the rapid onset of irregular heartbeat or a precipitous drop in blood pressure caused by bleeding into the pericardium. They are relatively uncommon and are difficult to diagnose.

14.13 ATHEROSCLEROSIS OR ARTERIOSCLEROSIS

Atherosclerosis or **arteriosclerosis** refers to the simultaneous development of an **atheroma** (from the Greek *athere* [meal] and *oma* [tumour]) in an artery and the sclerosis of its wall. An atheroma is a hard yellow plaque that gradually builds up on the inside of medium-sized arteries. The plaque consists of a necrotic (dead) core rich in cholesterol, surrounded by fibrous tissue. Sclerosis (from the Greek *skleros*, hard) means an abnormal hardening or fibrosis, which is the formation of excess fibrous material within a tissue. Sufferers may experience a sudden heart attack or stroke, but this belies the fact that, in most cases, the arteries of the victims have gradually become blocked by atherosclerosis.

Initially, atheromatous plaques start at the site in an artery, where the smooth muscle layer has thickened and been infiltrated with fibrous connective tissue (fibrosis) when cholesterol and other lipids have been deposited, and may become calcified. This condition is commonly referred to as 'hardening of the arteries'. During the condition, the affected artery expands as the plaque becomes larger so as to allow a more or less normal flow of blood. However, as the plaque increases in size, this becomes less possible, the lumen of the artery becomes narrower and a ballooning of the arterial wall causes it to weaken. Also, there is more likelihood of an embolus becoming trapped in the narrowed artery making the blockage worse. Healthy arteries are lined with endothelial cells, but the rough lining of a plaque-damaged artery seems to encourage the adhesion of platelets, which means they are common sites for the formation of a clot (thrombus).

The progression of atheromatous disease means the arteries become increasingly occluded and the threat of a heart attack or stroke increases. Some patients may receive warning in the form of chest pains if, for example, a coronary artery is partially blocked. The condition known as angina pectoris (see below) is a signal that the heart is not receiving sufficient oxygen. This is most likely to occur when the heart is working hard because of physical or emotional stress. However, for many people, there are no symptoms and they are completely unaware of their condition until the catastrophic event occurs. Some individuals have an inherited tendency to develop hypertension (*Section 14.17*), which promotes atherosclerosis and increases the risk of heart attack (*Section 14.14*) or stroke (*Section 14.16*) and can cause chronic damage to the endothelium lining the arteries, promoting atherosclerosis.

Angina

Angina pectoris is caused by myocardial ischaemia (*Section 14.9*). In males it presents as a crushing or squeezing pain in the chest and the discomfort may radiate into the neck, jaw, arms (especially the left) and sometimes the back. There may also be shortness of breath, abdominal pain, nausea and dizziness. The symptoms of angina in females can differ. For example, although chest pain is common, it may not be the only or most common symptom. Stabbing pains in the chest and abdomen can occur, together with discomfort in the neck, jaw, teeth or back, nausea and shortness of breath. Such differences may lead to delays in treatment.

Myocardial oxygen demand relates to the heart rate, left ventricular contractility and systolic wall stress. The demand for oxygen is increased by exercise, hypertension (*Section 14.17*) and left ventricular dilation, which may happen in chronic heart failure.

Stable angina occurs when atherosclerotic plaques block one or more of the coronary arteries. Under resting conditions, cardiac oxygen demand is quite low and is satisfied even by the diminished blood flow. However, when exertion or emotional stresses

increase this oxygen demand, ischaemia develops on the inner part of the myocardial wall. The response to exercise is variable: some patients may have excellent exercise tolerance one day and then develop angina with minimal exertion the next. In addition to causing pain, the ischaemia causes a decline in the output of creatine phosphate and ATP, and hence contractility is impaired. Stable angina is normally relieved by a short rest or by administering glyceryltrinitrate. The latter dilates the arteries, increasing blood, and therefore oxygen, supplies to the muscle, leading to less pain.

Ischaemic heart disease and stable angina can be distinguished from other conditions that cause chest pain on the basis of their characteristic symptoms and by a number of types of diagnostic tests, for example the ECG, exercise stress test and by using coronary angiography to obtain a direct radiographic visualisation, as described earlier (*Section 14.5*).

The management of angina is designed to control the symptoms and reduce any underlying risk factors. The drugs used include the nitrovasodilators, for example glyceryltrinitrate, β adrenoceptor blockers and calcium channel antagonists, as well as drugs that inhibit platelet aggregation and thrombosis (*Margin Note 14.4* and *Chapter 13*). In the case of stable angina, mortality is 2–4% a year if only one coronary artery is diseased, but increases with the number of diseased arteries.

Unstable angina is a dangerous clinical condition, often heralding an impending myocardial infarction (*Section 14.14*). In general, its symptoms resemble those of stable angina but are more intense and persistent, often lasting 30 minutes, and the pain is often resistant to glyceryltrinitrate treatment. The attacks may be frequent, becoming progressively more severe and prolonged, may be brought on by minimal exertion (or even during sleep) or may occur several days after a myocardial infarction. The episodes are preceded by a fall in coronary blood flow, which is thought to be the result of the periodic development of coronary thrombosis and vasoconstriction. These are triggered by coronary arterial disease. The thrombosis may be promoted by the turbulent blood flow associated with atherosclerotic plaques: there may also be damage to the endothelial lining of the blood vessels.

An ECG is taken to help in the diagnosis but, in addition, serum concentrations of C-reactive protein (*Chapter 4*) and amyloid A protein may be increased; these are classic markers of inflammation. Unstable angina is a medical emergency and treatment usually begins with aggressive drug therapy to control the symptoms and prevent further episodes and to try to reverse coronary vasospasm. Platelet glycoproteins IIIA and IIb are used in unstable angina to stabilise the clot in the narrowed coronary artery, which is causing the pain, prior to angiography and possibly angioplasty and stenting. Angioplasty is a procedure similar to angiography (*Section 14.5*) but the catheter delivers a small inflatable balloon to the narrowed portion of the coronary artery. When the balloon is inflated it opens the restricted section of the artery. Stents are small, coated spring-like structures that are deployed by cardiac catheters into the narrowed section following angioplasty. They act like miniature struts to maintain the opening of the vessel (*Figure 14.21*).

Urgent revascularisation via a heart bypass operation (see *Figure 14.29*) needs to be considered for patients at high risk or who are unsuitable for angioplasty/stenting due to significant coronary arterial disease. Treatment may include low doses of aspirin and the antiplatelet drug clopidogrel.

Variant angina is an intensely painful, transient spasm caused by a blockage of one of the coronary arteries. It is relatively uncommon, but can occur at rest. It is exacerbated by smoking and by cocaine use. About one-third of patients show no evidence of atherosclerotic lesions.

Cholesterol and lipoproteins

A number of nongenetic factors correlate with an increased risk of cardiovascular disease. Smoking, lack of exercise and a diet containing too much of and inappropriate fats (*Chapter 10*) all lead to an abnormal concentration of cholesterol in the blood. There are also inherited factors that predispose some individuals to develop hypertension (*Section 14.17*). Homocysteine (*Figure 14.22*) blood concentration is known to be increased in certain genetic conditions and coronary heart disease and is thought to have a role in increasing cholesterol levels. Homocysteine concentrations can be lowered by treatments with folate and cobalamin (vitamin B_{12}) with a subsequent reduction in blood cholesterol.

Cholesterol is an essential constituent of plasma membranes and is used in the biosynthesis of steroid hormones and bile salts (*Chapters 7* and *11*). However, atherosclerosis is largely due to problems with cholesterol. The body can synthesise it, but it is also obtained from the diet. Cholesterol is practically insoluble in plasma and, like

Margin Note 14.4
Thrombolysis

Thrombolysis is the dissolution of a blood clot blocking an artery. Fibrinolysis occurs when the inactive zymogen, plasminogen, is converted to the fibrin-dissolving active enzyme plasmin by endothelium-derived tissue plasminogen activator (t-PA). Plasmin also inactivates fibrinogen and coagulation factors V and VIII (*Chapter 13*).

Streptokinase is a bacterial protein that binds to a molecule of plasminogen. The resulting complex cleaves other molecules of plasminogen to produce more plasmin (*Figure 14.20*), which dissolves the fibrin. It can cause haemorrhage and, also, it can only be used once, since it may cause the patient to produce antibodies with the danger of an allergic reaction. Tissue plasminogen activator is produced commercially by recombinant DNA methods. It binds to fibrin and this has a greater effect on clot-associated plasminogen than on plasma plasminogen. It has the advantage that it is cleared from the plasma in a few minutes and is nonantigenic. Urokinase (u-PA) is also an endogenous plasminogen activator with properties similar to those of t-PA.

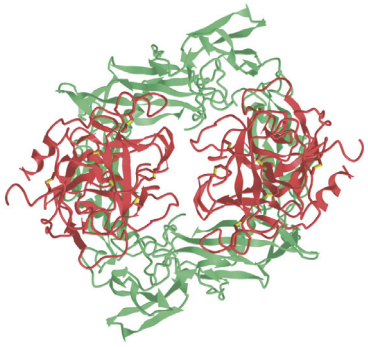


Figure 14.20 Molecular model of streptokinase–plasmin complex shown in green and red, respectively. PDB file 1BML.

Figure 14.21 Metal stent used in endovascular surgery. Shutterstock.com/Alexey Kamenskiy.

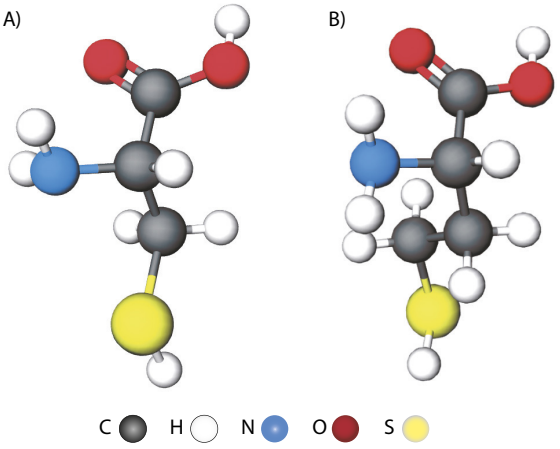


Figure 14.22 Computer drawn models of (A) cysteine and (B) homocysteine.

triacylglycerols, is transported in the blood in **lipoprotein particles**, which consist of a core of triacylglycerols and cholesterol esters surrounded by phospholipids, proteins and free cholesterol (*Figure 14.23*). Lipoprotein particles are classified by their densities: the greater the proportion of triacylglycerol, the lower the particle's density (*Table 14.1*). Several different types of lipoproteins occur in lipoprotein particles, called apolipoprotein-A and -B, -C, -D and -E, usually abbreviated to apoA, apoB, apoC, apoD and apoE.

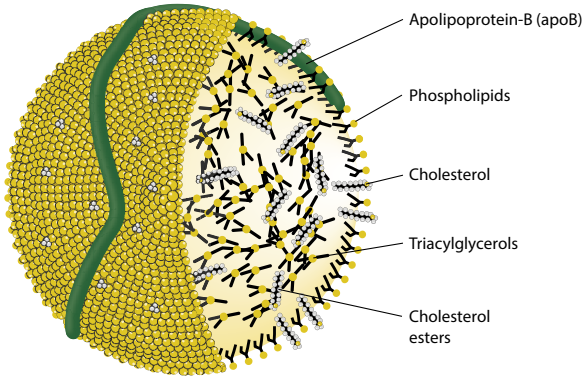


Figure 14.23 Schematic of a generalised lipoprotein particle.

	Chylomicrons	Very low-density lipoprotein	Low-density lipoprotein	High-density lipoprotein
Density (g/mL)	0.93	<1.006	1.019–1.063	1.063–1.21
Protein (%)	2	9	20	45
Phospholipid (%)	8	20	20	25
Free cholesterol (%)	1	7	10	5
Triacylglycerols (%)	85	55	10	8
Cholesteryl esters (%)	2	10	35	15
Apolipoproteins	B-48, C, E	A, B-100, C, E	B-100	A, C, D, E

Table 14.1 Composition of lipoprotein particles.

Most cells can synthesise cholesterol, but the bulk of biosynthesis occurs in the liver and intestinal enterocytes. Its subsequent transport around the body is a complicated process (*Figure 14.24*). Some cholesterol occurs in the chylomicrons (*Chapter 11*) but the liver exports this, together with the cholesterol it has synthesised and triacylglycerols

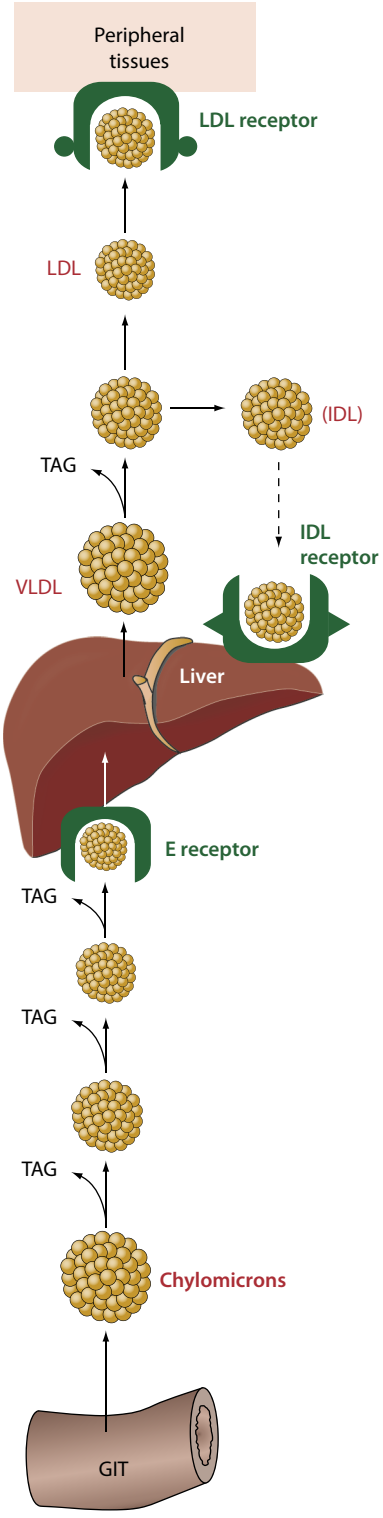


Figure 14.24 Schematic illustrating the transport of cholesterol. GIT, gastrointestinal tract; IDL, intermediate-density lipoprotein; LDL, low-density lipoprotein; TAG, triacylglycerols; VLDL, very low-density lipoprotein.

formed from dietary carbohydrate in the form of very low-density lipoproteins (VLDLs). These contain apolipoprotein B-100. Their major lipids are triacylglycerols, with less than 20% free cholesterol and cholesteryl esters present (*Table 14.1*). A primary determinant of the amount of VLDL secreted is the quantities of free fatty acids entering the liver. Further, if saturated fatty acids predominate, the VLDL particles are smaller but more numerous than if polyunsaturated fatty acids are in excess. A high carbohydrate diet also substantially increases the concentration of plasma VLDL particles. Newly secreted VLDL particles undergo a series of changes in the plasma. They acquire apoC and apoE proteins from high-density lipoproteins (HDLs; see below). The catalysed hydrolysis of triacylglycerols by lipoprotein lipase (*Figure 14.25*) of the endothelial cells of capillaries allows the fatty acids and glycerol to be taken up by tissues. Reducing the triacylglycerol content increases the density of the particles to form low-density lipoproteins (LDLs), which are the principal carriers of cholesterol in the plasma. Low density lipoprotein particles serve as the major source of cholesterol for most tissues.

The concentration of LDL in the plasma correlates positively with the incidence of coronary heart disease. Hence, LDL is often referred to as 'bad' cholesterol. However, it is only 'bad' in excess and when it is combined with other risk factors. High-density lipoproteins are secreted by the liver, but they are also formed by modifications to chylomicrons and VLDL. High-density lipoproteins (*Table 14.1*) can pick up cholesterol from tissues, which is essentially the opposite function to that of LDLs. The particles probably acquire cholesterol from the cell surface membranes and convert it enzymatically to cholesteryl esters. Consequently, HDL may be considered 'good' cholesterol.

Familial hypercholesterolaemia

Familial hypercholesterolaemia is an autosomal dominant condition that affects heterozygotes, with a frequency of 1 in 500. In health, circulating LDL particles are recognised by protein receptors on the surfaces of liver cells and are removed from the plasma by endocytosis into the hepatocytes. Here, their various components are metabolised, stored or recycled. Several forms of familial hypercholesterolaemia have been described that are associated with mutations of the genes encoding apoB or receptors of LDL particles, which are found on the surfaces of liver cells. Sufferers of familial hypercholesterolaemia present with whitish-yellow deposits largely of cholesterol on the tendons called **xanthomas**, with opaque fatty deposits around the periphery of the cornea called **corneal arcuses**, and the onset of coronary heart disease before the age of 10 years in homozygotes.

It is necessary to know the concentration of LDL cholesterol in serum to assess the risks to a person, diagnose an ailment and monitor the progress of a patient. The blood cholesterol concentration in homozygotes is 15–30 mmol/L, compared with normal values of approximately 5.0 (*Table 14.2*).

Lipid component	Concentration (mmol/L)
Total cholesterol	<5.2
Low-density lipoprotein cholesterol	<3.6
High-density lipoprotein cholesterol	
Female	>1.2
Male	>0.9
Triacylglycerols	0.4–1.8

Table 14.2 Desirable concentrations of cholesterol, lipoproteins and triacylglycerols in adults.

Hypercholesterolaemia usually shows no symptoms. However, if left untreated, it can lead to an increased risk of heart disease or stroke and is therefore a hidden risk factor. Indeed, high blood cholesterol causes a third of all ischaemic heart disease. In England, it leads to over 7% of all deaths and affects up to 60% of adults, hence the necessity of determining the concentration of cholesterol in blood.

The most common drugs used in its treatment are statins, such as atorvastatin and simvastatin. Statins are inhibitors of 3-hydroxy-3-methylglutaryl-coenzyme A reductase (HMG-CoA reductase), which catalyses the NADPH-dependent reduction of HMG-CoA to mevalonate, a key step in the biosynthesis of cholesterol (*Figure 14.26*). They are

Figure 14.25 Molecular model of the endothelial lipoprotein lipase. Its bound glycans and Ca^{2+} are shown in brown and as a red sphere, respectively. The lipoprotein that binds the enzyme to the plasma membrane is shown in red. PDB file 6E7K.

Margin Note 14.5
Tangier disease

The removal of cholesterol from cells uses an ATP-binding cassette (ABC) transporter that hydrolyses ATP in an active transport process (see also *Section 16.2*) encoded by *ABCA1* sited at 9q31.1. The rare autosomal recessive disease called Tangier disease is associated with mutations in the gene encoding this transporter and leads to an accumulation of cholesterol esters in the tissues and an almost complete absence of cholesterol in HDL particles. The name 'Tangier' comes from an island off the coast of Virginia, USA, where inbreeding amongst the families living there in isolation made it relatively common. The outlook for sufferers of this mutation is premature coronary heart disease.

some of the most widely prescribed medicines because of their efficacy in reducing cardiovascular events and mortality.

In the UK, prophylactic daily treatment with a low dose of a statin was recommended by the National Institute for Health and Care Excellence (NICE) for those whose clinical profile suggested they had a 10% or greater risk of a cardiovascular event within 10 years. However, given that the adverse side effects of muscular pains associated with statins appear to have been over estimated, in 2023 NICE published draft updated guidelines recommending that statins should also be offered to people at the lower threshold of a 5% greater risk of cardiovascular events within 10 years. It is thought that the use of statins will prevent 70 in 1000 people from having heart disease or a stroke over a 10-year period.

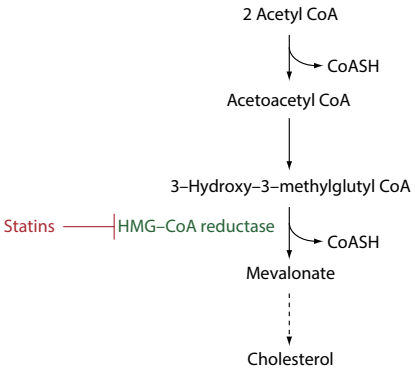


Figure 14.26 Outline of the biosynthesis of cholesterol, emphasising the role of statins in inhibiting 3-hydroxy-3-methylglutaryl-coenzyme A reductase (HMG-CoA reductase), which catalyses the key reaction of HMG-CoA to mevalonate. CoASH, coenzyme A (the SH is used as a reference to its 'active' sulphydral group).

14.14 MYOCARDIAL INFARCTIONS

Myocardial infarctions, 'heart attacks' (*Marginal Note 14.2*), are the second most common cause of death in the UK but, surprisingly, were hardly known before 1910. An **infarction** is the death of a section of tissue because its blood supply has been cut off; an **infarct** is the segment of tissue affected. In general, if one artery is blocked, neighbouring arteries with communicating branches can compensate and tissue death is limited. Infarction occurs in places where small arteries do not communicate with one another, such as in the kidney; where all the arteries together supply only enough blood for the whole organ, such as in the brain; or where alternative arteries are also blocked and cannot take over. The latter is what happens to the coronary vessels in many middle-aged hearts, particularly in the developed world. Myocardial infarction almost always occurs in patients with atheroma in the coronary arteries resulting from sudden coronary thrombosis, usually at the site of a fissure or rupture of the surface of an atheromatous plaque. There may be haemorrhage into the plaque with local coronary spasms. Irreparable damage can begin after only 20 minutes of occlusion. After about six hours, the site of infarction of the myocardium is pale and swollen and after 24 hours necrotic tissue appears deep red owing to the haemorrhage. Subsequently, an inflammatory reaction develops, and the infarcted tissue turns grey in colour.

Patients present with severe intermittent chest pain that is similar in character to the angina that can occur on exertion, but usually presents at rest and lasts several hours. Sometimes, however, the pain is less severe and may be mistaken for indigestion. The episodes of pain may become more frequent, but about 20% of patients have no pain. If there is pain, the onset is usually, but not always, sudden. The patient may feel restless and there is often sweating, nausea and vomiting. The most recognisable pain is in the middle of the chest, which may spread to the back, jaw or left arm. The condition, once recognised, is a medical emergency. Half of the deaths occur in the first three to four hours after the symptoms begin, so the sooner treatment begins the better the chances of survival.

Diagnosing and treating myocardial infarctions

The diagnosis of a myocardial infarction is usually made on the basis of the clinical symptoms and ECG findings and is confirmed by the characteristic changes in serum biomarkers (*Box 14.5*). Biomarkers are biological molecules whose concentration in blood changes in response to a specific disease. Indeed, measurements of biomarkers can be used in diagnosing a disease or to monitor and predict its outcome following treatment. Biomarkers can be enzymes or other proteins, for example structural ones, molecules of intermediary metabolism and even cell signalling molecules. Most biomarkers that have been used to investigate myocardial infarctions have been enzymes.

Enzymes are largely intracellular and are released into the blood when cells are damaged or die. Enzyme activities that have been of greatest value in investigating myocardial infarctions are creatine kinase (CK), lactate dehydrogenase (LDH) and aspartate transaminase (AST, previously known as glutamate oxaloacetate transaminase [GOT]). Plasma activities of these enzyme are increased in about 95% of myocardial infarctions and sometimes increase to high values. The degree of increase gives a rough estimation of the size of the infarct but is of little prognostic value. A second and subsequent rise after returning to normal levels may indicate an extension of the damage. All three enzymes tend to show normal blood activities until at least four hours *after* the onset of chest pain due to the infarction, and so blood samples should not be taken until after this time. However, determining enzyme activities has several limitations. They only give a 'rough' guide to tissue damage, may be released due to changes in the functions of cells as opposed to their death or damage, and may lack specificity, that is, these enzymes may be found in tissues in addition to heart muscle. For these reasons, enzyme measurements are being phased out and there is considerable interest in using myoglobin

BOX 14.5 BIOMARKERS IN MYOCARDIAL INFARCTIONS

Biomarkers are measured in blood samples and are useful in diagnosis and monitoring myocardial infarctions (*Figure 14.27*). Aspartate aminotransferase (AST) activity increases in blood approximately 12 hours after a myocardial infarction and peaks at about 36 hours before returning to normal values approximately 60 hours after the event. However, AST is also a marker for damage to other organs, such as the liver, and so lacks specificity. Lactate dehydrogenase (LDH) activity in the blood increases later than AST, that is, 36 hours after a myocardial infarction, and is detectable for up to 72 hours, and so was used as a late marker of an attack. However, LDH activity is widespread in the body, so this marker also lacks specificity. Both these enzyme biomarkers are now redundant. Creatine kinase (CK) activity in the blood increases four to six hours after a myocardial infarction and peaks at 24 hours. However, CK is found in all muscles of the body and so its release can lack specificity. There are three isoenzymes of CK:

1. CKMM, found mainly in skeletal muscle
2. creatine kinase myocardial band (CKMB), found mainly in cardiac muscle
3. CKBB, found in brain and smooth muscle

Analysing CKMB activity is the most sensitive of the three for diagnosis of myocardial infarction. The activity of CKMB in the blood increases slightly earlier than the overall CK value, and peaks 21 hours after a myocardial infarction. However, although cardiac muscle contains more of the CKMB isoenzyme than the other types, it also contains significant amounts of CKMM. Similarly, skeletal muscle contains mostly CKMM but also has a significant amount of CKMB. Thus, the increase in CKMB is not absolutely specific for cardiac damage. However, measurements of CKMB were the best test for a myocardial infarction for many years.

Myoglobin, a haem protein, is released two to three hours after a myocardial infarction, reaching a peak concentration in the blood after 12 hours. Unfortunately, myoglobin too is found in other muscle tissue, for example skeletal, and so like CK and CKMM is not a specific marker.

The amounts of troponin subunits increase in the blood within a few hours of a myocardial infarction and remain elevated for one to two weeks, making both troponin C and I suitable for diagnosing early and late myocardial infarctions. Diagnostic sensitivity of troponin reaches 100% 12 hours after onset of symptoms; hence, a myocardial infarction can be excluded if a troponin test on a blood sample collected 12 or more hours after onset of chest pain gives a negative (undetectable) result. Furthermore, any increase of troponin to a clinically significant level indicates a greater risk of a cardiac event occurring over the next 0–60 days.

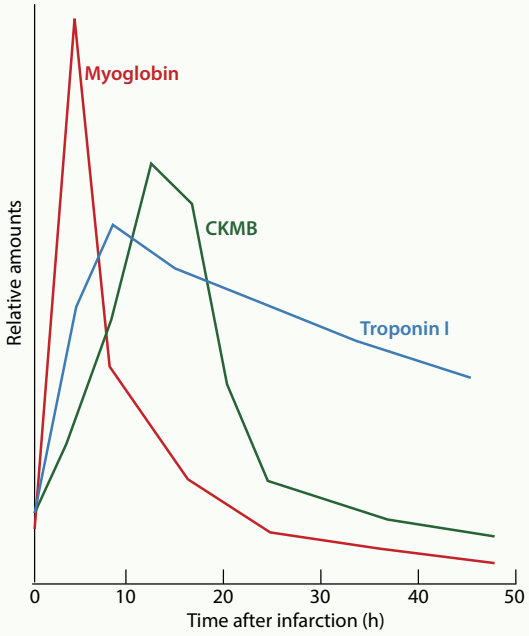


Figure 14.27 Graph illustrating the changes in concentrations of selected biomarkers following a myocardial infarction (heart attack). Redrawn from Aus Nurse Educator (Aus NE). CKMB, creatine kinase myocardial band (that is, the form specific to cardiac muscle).

and troponin as biomarkers for myocardial infarctions. Myoglobin is a haem-containing protein that strongly binds oxygen (*Chapter 13*). It is found in the cytoplasm of cardiac muscle cells and is released within two to three hours after a myocardial infarction, much earlier than enzyme biomarkers, so there is interest in its use as an early indicator of myocardial infarction. However, myoglobin too lacks absolute specificity, as it is found in other cell types, such as those of skeletal muscle.

The current biomarker of choice is troponin (*Box 14.5*), which is used as a marker for both early and late myocardial infarction. The troponin–tropomyosin complex (*Figure 14.28*) of cardiac muscle fibres is part of the contractile apparatus in these cells. Troponin is a heterotrimer consisting of three subunits: troponin T (function: tropomyosin binding), troponin I (inhibits muscular contraction) and troponin C (calcium binding). The structures of troponin T and I in cardiac and skeletal muscles differ, but it is possible to unambiguously identify the cardiac troponin T (cTnT) and cardiac troponin I (cTnI), giving their measurements a high degree of specificity. During a myocardial infarction, cardiac muscle cells die and then lyse, releasing the troponin subunits into the blood. Therefore, increased concentrations of cTnT and cTnI in blood samples indicate myocardial necrosis, as they are cardiospecific (*Figure 14.27*).

Treating cardiac infarctions usually involves giving the patient aspirin, which should improve the chances of survival by reducing the clot in a coronary artery. A β blocker may also be given to slow the heart rate and reduce its workload, and oxygen may be

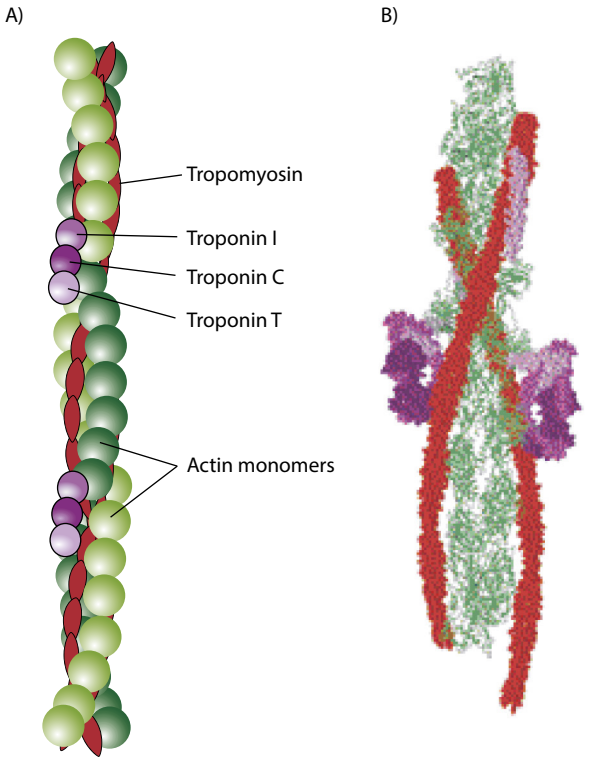


Figure 14.28 (A) Schematic of a cardiac muscle complex. (B) Molecular model of a cardiac muscle complex. Actin monomers are coloured green, tropomyosin complexes are red and troponins C, I and T are depicted in increasingly lighter shades of magenta. PDB file 6KN8.

given through a face mask to increase its supply to the heart. Blood clots in an artery can often be cleared by intravenous thrombolytic therapy (*Box 14.3* and *Margin Note 14.4*). The indication for thrombolytic treatment is usually based on the clinical presentation and the ECG picture rather than on the activities of plasma enzymes. Treatment must be given within six hours of the start of the heart attack to be effective. After six hours, it is likely that some of the damage will be permanent, and the patient could be compromised and some may die. Most patients who survive for a few days after an attack can expect a full recovery, but about 10% will die within a year. The majority of deaths occur in the subsequent three to four months in patients who continue to have angina, arrhythmias and subsequent heart failure.

Coronary bypass surgery is a possible treatment for patients who have angina and coronary arterial disease that is not too widespread. It improves exercise tolerance, reduces symptoms and decreases the number of drugs that are needed. Heart bypass surgery involves grafting a length of healthy 'surplus' blood vessel, such as the saphenous vein from the leg or left internal mammary artery, between the aorta and the coronary arteries distal to any stenosis (narrowing) to take blood from the aorta past the obstructed region. Such grafts work well for 10 years or more. A bypass improves survival in patients with severe atherosclerotic disease in all their major coronary arteries (*Figure 14.29*).

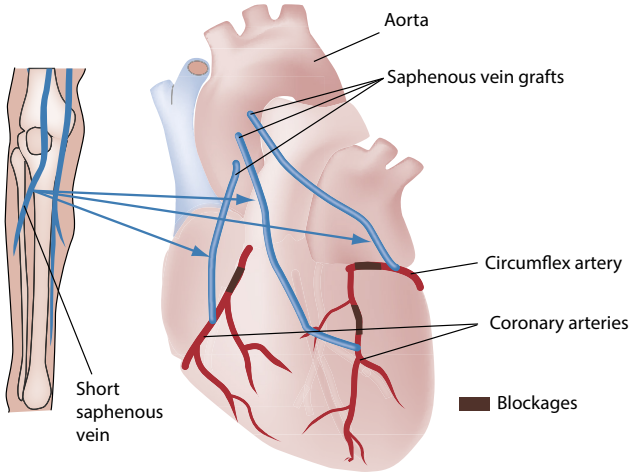


Figure 14.29 Schematic overview of bypass surgery.

14.15 PERICARDIAL DISEASE

Inflammation of the pericardium results in an acute or chronic **pericarditis**. Acute pericarditis has a sudden and often painful onset. The inflammation causes fluid and blood components, such as fibrin, erythrocytes and leukocytes, to pour into the pericardial space. Acute pericarditis may be caused by a viral infection, in which case the condition may be painful but short lived with no lasting effects or result from a number of other causes, for example cancer, heart attack, acquired immunodeficiency syndrome (AIDS), kidney failure, heart surgery and the side effects of certain drugs, some of which are life threatening.

Symptoms of acute pericarditis include fever and chest pain that typically extends to the left shoulder and down the left arm. There are also characteristic heart sounds. The treatment for acute pericarditis is to hospitalise the patient and treat with anti-inflammatory drugs, such as aspirin or ibuprofen, that also reduce the pain. Further treatment depends on the underlying cause. Individuals with cancer that has invaded the pericardium rarely survive longer than 12–18 months. Given that the pericardium is not absolutely essential to life, it can be removed without significantly affecting the functioning of the heart.

The chronic form of the disease develops gradually and is long-lasting. Usually the cause is unknown, but cancer and reduced thyroid function have been implicated.

14.16 DISORDERS OF ARTERIES AND VEINS

A number of clinical disorders are associated with arteries and veins. These include:

- peripheral arterial disease
- arterial aneurysms
- aortic dissection
- strokes
- varicose veins
- deep vein thromboses

Peripheral arterial disease

Peripheral arterial disease can affect the abdominal aorta and its major branches, including the arteries to the legs. Obstruction can be sudden or gradual. Most patients with peripheral arterial disease have atherosclerosis that gradually narrows the arteries. Partial occlusion can also result from a blood clot resulting in a sudden decrease in the oxygen supply. A sudden, complete obstruction normally results from a clot lodging in a narrowed artery. Emergency surgery or the use of thrombolytic drugs may be needed to remove the obstruction.

Aneurysm

An **aneurysm** is a round or tube-like bulge that usually develops in mechanically weak areas of an arterial wall. There are many reasons for the development of aneurysms, but high blood pressure and cigarette smoking increase the risks. Also, a blood clot may form in the aneurysm. If an aneurysm occurs in the aorta, an aortic aneurysm, rupture, haemorrhage and separation of the layers of the wall (called **dissection**; see below) can occur, with disastrous results. Such conditions can be immediately fatal, although most take years to develop. Aortic aneurysm can develop anywhere along its length but over 70% of them occur in the segment that runs through the abdomen. An individual with abdominal aortic aneurysm usually becomes aware of a pulsing sensation in the abdomen, with a deep penetrating pain mainly in the back. The aneurysm may rupture, with severe internal bleeding, and the patient will typically go into severe shock (*Section 14.17*). Aortic rupture is frequently fatal. The treatment of ruptured aneurysms involves surgical repair, which is extremely risky. However, many patients are diagnosed early during a routine examination and then surgery to insert a synthetic graft can be used to repair the aneurysm, with a good chance of success.

Aortic dissection

In an aortic dissection, the inner lining of the vessel wall tears, allowing blood to surge through the tear, splitting the middle layer and creating a new channel. The condition usually results from a deterioration of the arterial wall caused, in most cases, by high blood pressure, but it can also be a consequence of certain hereditary conditions. The clinical picture includes stroke (see below), heart attack, sudden abdominal pain, nerve

Margin Note 14.6
Abdominal aortic aneurysm screening

The health authorities of a number of countries operate screening programmes for the early detection of abdominal aortic aneurysms (AAAs). In the UK, the National Health Service offers an AAA screening programme to all men over 64 years old. The test is not offered to anyone with a low risk of an AAA, which includes men under 65 years old, women and patients who have already been treated for an AAA. The screening test uses an ultrasound scan of the abdomen to assess the presence and size of any aneurysm in the abdominal aorta, which has the clinical risk of a potentially fatal aortic dissection. If an aneurysm is not found, then no further scans or action are required. However, an aneurysm smaller than 30 to 44 mm requires further yearly scans to assess any increase in its size. Medium sized aneurysms of 45 to 54 mm require assessment scans every three months, with accompanying advice from a specialist nurse. Those men unfortunate to have large aneurysms of 55 mm or more are referred to a vascular surgeon for additional tests and to discuss possible treatments.

damage and an inability to move a limb. Treatment in intensive care is required. Drugs, for example β-blockers such as propranolol (*Figure 14.14*), are given to reduce the heart rate and lower the blood pressure. Thereafter, a decision needs to be taken as to whether to carry out surgery to replace the portion of damaged blood vessel with a synthetic graft. Untreated, about 75% of patients die within two weeks, but, following treatment, several years of life are possible in the majority of cases, even though the death rate from surgery is 15% or more.

Stroke

If the blood supply to the brain is disrupted for any length of time, brain cells can be permanently damaged or die due to the lack of oxygen and this is called a **cerebrovascular accident** (CVA) or **stroke** (*Figure 14.30*). Brain cells are also damaged if bleeding into the brain occurs. Therefore, strokes can be either ischaemic or haemorrhagic. In ischaemic stroke, the blood supply to part of the brain is cut off because of either atherosclerosis or a clot blocking a blood vessel. In haemorrhagic stroke, a blood vessel bursts, preventing normal flow of blood and allowing it to leak into an area of the brain and destroy it.

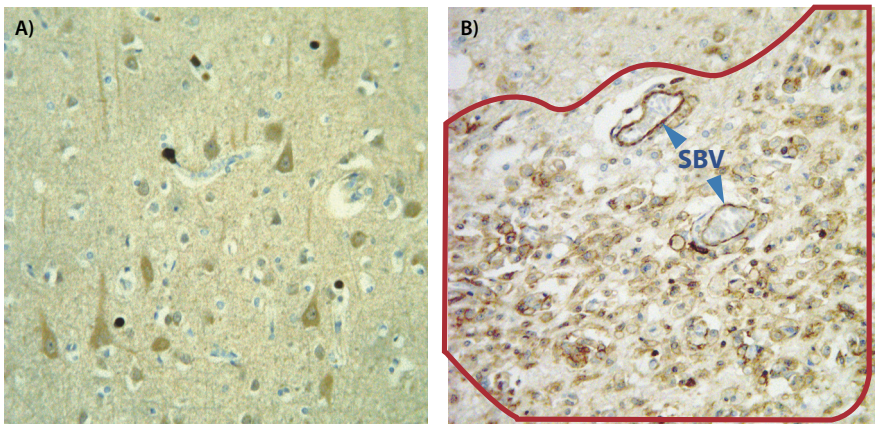


Figure 14.30 Photomicrographs showing (A) normal brain tissue with the nuclei of numerous neurons prominent. (B) Brain tissue following an ischaemic stroke. The area of damaged (infarcted) tissue is enclosed. Note the surviving blood vessels (SBVs). ⏎

In the developed world, strokes are the most common cause of disabling neurological damage resulting, typically, in a loss of speech and/or loss of motor function on one side of the body. High blood pressure and atherosclerosis are the major risk factors. The incidence is decreasing because the importance of controlling high blood pressure and dealing with inappropriate blood cholesterol concentrations has been recognised. Clinicians can usually diagnose a stroke from the clinical history of events and a physical examination. Computed tomography scans and MRI (*Chapter 18*) are used for differential diagnoses. Many people who have had a stroke recover some or all of their normal functions, but others may be mentally and physically devastated, unable to speak or move normally. However, over 65% of stroke patients survive three or more years. Factors that affect survival include the patient's age and overall health and the severity of the stroke. The younger the patient, the better the prognosis.

Since each area of the brain is served by specific blood vessels, the area where the cerebrovascular accident occurs determines which part of the body becomes disabled. The loss of function is greatest just after the stroke. Some function may return subsequently because, although some brain cells die, others may recover and take over a given function, at least partially. The immediate treatment is to give oxygen to try to prevent further damage. Anticoagulants may be given if the stroke is ischaemic, but these are of little use once the stroke is completed. Surgery is not normally of much use, since the brain cells are already dead. For the survivors, a long period of rehabilitation may be required.

Varicose veins

Varicose veins are enlarged superficial veins, which occur particularly in the legs (*Figure 14.31*). The cause(s) of this condition is(are) not precisely known, but it is probably due to a weakness in the walls of the veins, which may be inherited. Over time, the veins lose their elasticity and they stretch and become wider. They may take on a tortuous,

Figure 14.31 Picture showing external appearance of varicose veins of the long saphenous system of the thigh and calf. Courtesy of J. Guy, The Royal College of Surgeons of Edinburgh, Scotland. ⏎

snake-like appearance and cause bulges in the skin over them. The widening causes the valve cusps to separate (*Figure 14.32*) and, as a result, the veins fill with blood when the person stands and the veins bulge even more. As well as being unsightly, the legs feel tired and the veins ache, and after removing socks or stockings the legs may itch. Other complications include bleeding and blood clots. The latter can lead to deep vein thrombosis (see below), and minor injuries may cause an ulcer that fails to heal. Varicose veins are common during pregnancy but these usually improve during the two to three weeks following delivery. Hypertension and obesity may have parts to play in the propensity for varicose veins.

Treatment for varicose veins

Varicose veins cannot be cured. The symptoms may be relieved by elevating the legs on a stool when sitting down and wearing elasticated stockings to compress the veins and prevent them hurting. Itching may be eased by applying calamine lotion or paraffin wax-based moisturisers to the affected areas. Varicose veins may be treated surgically by either stripping or injection therapy. Stripping involves removing as many of the varicose veins as possible. The superficial veins play a less important role than the deep veins in returning blood to the heart and their removal does not impair the circulation significantly. Two incisions are made: one in the groin and one at the ankle while the patient is under general anaesthetic. The saphenous vein (*Figure 14.33*) is then removed by threading a flexible wire through the vein, which is pulled to remove the entire vein. However, surgery does not remove the tendency to develop new varicose veins. Injection therapy involves sealing the veins by injecting an irritant solution that causes a thrombus to form so that no blood flows. Healing of the thrombus causes scar tissue, which can block the vein. However, the thrombus may dissolve, allowing the vein to reopen. Injection therapy was popular but has fallen into disrepute, probably because of poor techniques with resulting complications. The more modern techniques, if carried out carefully, appear more successful.

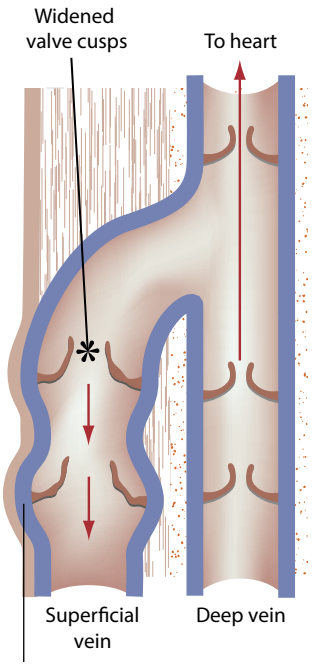


Figure 14.32 Schematic to show the widening of the valve cusps in varicose veins.

Deep vein thrombosis

Deep vein thrombosis, as the name implies, is clotting of the blood in a deep vein. Like varicose veins, they primarily affect the legs. They are potentially dangerous. All or part of the clot or thrombus can break loose and lodge in a narrow artery, with painful results. In the lungs, obstructions to the flow of blood causes a pulmonary embolism. In serious cases, this can result in blockage of all or nearly all of the blood travelling from the right side of the heart to the lungs, rapidly causing death. Such serious consequences are not common, but it is impossible to predict what will happen once a thrombus has formed. It is believed that the possible causes of this condition are an increased tendency of the blood to clot, which can happen with some cancers, for example liver cancer (*Chapter 17*), and, very occasionally, with the extended use of progesterone oral contraceptives (*Chapter 7*). Slowing the movement of blood in the veins may also occur in prolonged bed rest, sitting on a long flight and some types of injury or major surgery. The condition is difficult to diagnose until the thrombus moves. Deep vein thrombosis may be prevented to some extent by flexing and extending the ankles from time to time and by wearing elastic stockings. Anticoagulant therapy may also be appropriate in some cases.

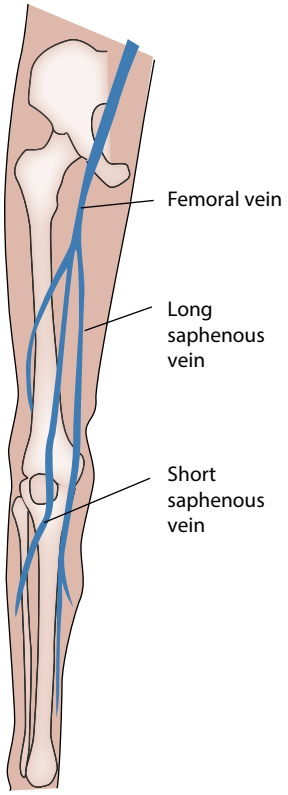


Figure 14.33 Schematic showing the saphenous veins of the leg.

14.17 CLINICAL PROBLEMS ASSOCIATED WITH BLOOD PRESSURE

The blood pressure (*Section 14.4*) of any individual varies with their age, sex, physical activity and emotional state. Blood pressure also varies with ethnic origin, for example African Americans are much more likely to have high blood pressure than white populations. The normal upper limit at age 20 years is 150/90 mmHg (150/12 kPa), which can increase to 160/95 (21.3/12.7) and 170/95 (22.7/12.7) at 40 and 60 years, respectively. If these pressures are found consistently then they are abnormally high.

Three major factors determine the blood pressure:

1. the amount of blood pumped by the heart
2. the volume of blood in the blood vessels
3. the elasticity of the blood vessels

Thus, the more blood pumped by the heart, the higher the pressure. If the heart beats more slowly or its contractions are weakened, such as may happen after myocardial infarction (*Section 14.14*), then less blood will be pumped. Equally, a rapid heartbeat can result in inefficient pumping. The greater the volume of blood in circulation, the more

likely the blood pressure is to be high. Conversely, a loss of blood by bleeding or as a result of dehydration will have the opposite effect. The smaller the capacity of the blood vessels, the higher will be the resulting blood pressure. Consequently, a dilation of the blood vessels will lower the blood pressure.

Sense organs in the neck and chest continually monitor blood pressure and trigger physiological responses if it changes. The actions taken might be to modify and strengthen the heartbeat; to regulate the kidneys to alter the amount of water excreted, which, in turn, changes the volume of blood circulating; or to constrict or dilate the blood vessels. However, these compensatory mechanisms have limits; for example, if too much blood is lost as a result of bleeding, there is little that can be done. Action must be taken to stop the bleeding (*Chapter 13*) and to transfuse blood or fluid to make up the volume (*Chapter 6*). Furthermore, these compensatory mechanisms may themselves fail in certain conditions.

Low blood pressure (hypotension)

Hypotension is an abnormally low blood pressure, below 100/60 mmHg (13.3/8 kPa), at any age. Too low a blood pressure can lead to dizziness and fainting. Fainting or **syncope** is the result of a temporary inadequacy of oxygen and nutrients and is usually associated with a temporary decrease in blood flow. This can happen in people with an abnormal heart rhythm when they suddenly begin to exercise or the heart rate is too slow. However, there are many other possible causes of fainting including anaemia, hypoglycaemia, **hypocapnia** (a lower than usual concentration of CO_2 in the blood) or hyperventilation. The latter may be caused by anxiety. Usually, lying flat is all that is needed for the individual to regain consciousness, but checks should be performed to eliminate more serious conditions and this may mean, for example, an ECG examination.

Shock

Shock occurs when the blood pressure falls too low to sustain life following the collapse of the cardiovascular system. This is a severe insult to body cells, since, when deprived of oxygen and nutrients for any length of time, they quickly become irreversibly damaged and die. Shock may result from low blood volume, inadequate pumping by the heart or excessive vasodilation as can occur in extreme allergic reactions (*Chapter 5*). Low blood volume may be the result of bleeding following serious trauma. Inadequate pumping by the heart may be the result of a heart attack, pulmonary embolism, the failure of a heart valve, an irregular heartbeat or drug toxicity. Head injuries, liver failure, poisoning, severe bacterial infections and drug reactions may all lead to excessive blood vessel dilation.

Unless treated promptly, shock is usually fatal. When shock results from a sudden loss of blood, for example in an accident or a haemorrhage, the first person on the scene should aim to stop the bleeding, keep the victim warm and raise the legs slightly to improve the return of blood to the heart. Emergency personnel may provide mechanically assisted breathing if it has stopped and fluid or blood to increase the blood volume. Other treatments will depend on the cause of the shock.

High blood pressure (hypertension)

Hypertension occurs when the blood pressure is higher than that which is regarded as normal, regardless of its cause: systolic pressure at rest of 140 mmHg (18.7 kPa) or greater and a diastolic pressure of 90 mmHg (12 kPa) or greater. In fact, both are usually elevated in hypertension. It does not usually cause symptoms, at least for many years, and often tends to go undetected unless the blood pressure happens to be measured for some other reason. Nevertheless, it is sometimes referred to as a 'silent killer' because there tend to be no symptoms until some vital organ is damaged. Mortality and morbidity rise continuously with increasing blood pressure. However, the risk is not linear and rises more steeply at higher pressures.

Initially, an abnormal heart sound, indicating hypertension, may be detected using a stethoscope. A diagnosis of hypertension can be made on the basis of an elevated blood pressure reading of 140/90 mmHg or more when measured several times. A single reading on a given day is unreliable. Obviously the higher the values, the more serious the condition must be considered. It is possible to judge the seriousness of the condition by examining the arterioles at the back of the eye to determine the degree of damage to the retina, as hypertension can cause retinopathy. In addition, ECG and echocardiography can detect an enlargement of the heart brought on by the increased workload. Kidney damage may be detected by urine analyses.

The cause of hypertension can be identified in less than 10% of patients. These are usually kidney disease, a hormonal disorder or the use of oral contraceptives. Thus, in

most cases, the primary cause cannot be identified, and this form is referred to as essential hypertension. Many factors are probably responsible. For example, in older people, the larger arteries lose their flexibility and become stiffer. Consequently, when the heart pumps blood, they cannot expand and the pressure increases. If the kidneys malfunction such that the urinary output is decreased, more fluid will be retained. Obesity, stress, a sedentary lifestyle, excessive amounts of alcohol and too much salt in the diet can also contribute.

Untreated hypertension increases the chances of a person developing heart diseases, such as cardiac failure or myocardial infarction (*Section 14.14*), kidney failure or a stroke (*Section 14.16*). Stopping smoking, reducing weight, salt intake and cholesterol concentrations reduce the risk. In general, patients do not have to restrict their activities as long as their blood pressure is controlled. Various drugs are available as part of a treatment programme. These include diuretics to help the kidneys eliminate water and salt (*Chapter 8*), adrenergic (β) blockers (for example atenolol, bisoprolol or carvedilol [*Figure 14.34*]) to block the effects of the sympathetic nervous system, and ACE inhibitors (*Margin Note 14.3*), which lower the blood pressure by stimulating arterial dilation.

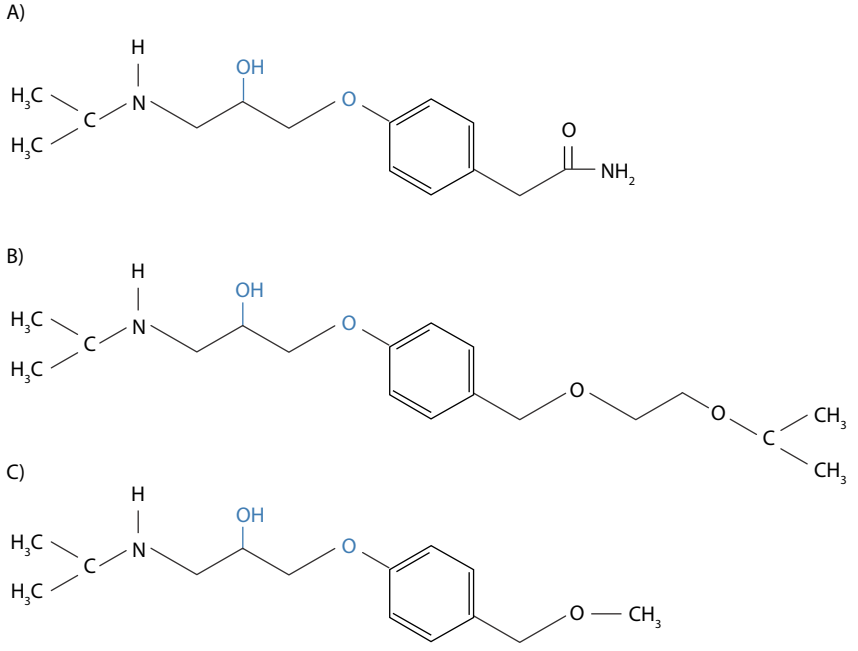


Figure 14.34 Structures of the adrenergic (β) blockers (A) atenolol, (B) bisoprolol and (C) metoprolol. See also *Figure 14.14.*

14.18 CASE STUDIES

CASE STUDY 14.1

Jim, a 52-year-old man, has suffered from angina over the past 18 months. He has also been suffering from lethargy, constipation and loss of concentration. His serum specimen was analysed and yielded the following results (reference ranges are given in parentheses):

Thyroid-stimulating hormone	>100 (0.35–4.1 mU/L)
Cholesterol	12.8 (<5.2 mmol/L)
Triacylglycerols	1.4 (0.5–0.9 mmol/L)

Question

(a) Explain these results.

CASE STUDY 14.2

Andrew, a recently retired lecturer, was admitted to hospital with chest pain that had developed in the evening after spending a day working in the garden. There were no signs of a myocardial infarction on his ECG. He was monitored at the hospital coronary unit for 12 hours. His pain subsided rapidly, after which he was discharged. His serum troponin T (on admission and after 12 hours) and CK (after 12 hours) measurements were:

	On admission	After 12 hours	Reference range
Troponin T	<17	<17	(<17 ng/L)
CK activity		970	(40–320 IU/L)

Question

(a) Explain these results.

CASE STUDY 14.3

Roger, a 37-year-old bank manager, has corneal arcus and xanthomata. He has normal blood pressure and weight and is a nonsmoker. However, he has a family history of myocardial infarction, as his father died following one at the age of 40 years. A blood specimen was taken and analysed for its lipid content and yielded the following results (reference ranges are given in parentheses):

Cholesterol	16.8	(<5.2 mmol/L)
Triacylglycerols	2.1	(0.4–1.8 mmol/L)
LDL cholesterol	14.2	(<3.6 mmol/L)
HDL cholesterol	1.3	(>1.2 mmol/L)

Questions

(a) How do you account for this abnormal lipid profile?
(b) How should Roger be treated?

CASE STUDY 14.4

Tom, a 59 year old, had a superficially healthy appearance but a family history of heart disease. He attended a routine NHS health check at his local surgery and was found to have a blood pressure of 206/90 mmHg (27.5/12 kPa). His assessment showed the following:

- heavy cigarette smoker (35/day)
- alcohol intake of 120 units/week
- body mass index: 24

Tom was referred to his general practitioner, who undertook some blood investigations and persuaded him to have home blood pressure monitoring. He returned to see the practice senior nurse and further investigations showed the following (reference ranges in parentheses):

- home blood pressure monitoring the average of eight readings: 180/95 mmHg (24/12.7 kPa)
- cholesterol: 7.2 (<5.2 mmol/L)
  - LDL cholesterol: 5.8 (<3.6 mmol/L)
  - HDL cholesterol: 1.4 (>1.2 mmol/L)
- normal liver function tests

Question

(a) What advice should be given to Tom?

14.19 SUMMARY

The cardiovascular system comprises the heart and the blood vessels that supply the organs and tissues of the body with, for example, oxygen and nutrients, and remove CO_2 and other waste. The heart pumps blood through the pulmonary circulation, which oxygenates the blood, and the systemic circulation around the rest of the body. The heart pumps nonstop for millions of beats during a person's lifetime. However, there are a number of conditions that can cause serious clinical problems. These can be investigated by listening to the heart sounds using a stethoscope, by measuring its electrical activities by an ECG, and by studying its function by echocardiography or CT. The rhythm of the heartbeat may be faulty, and this may require correction, including the insertion of an artificial pacemaker. Heart failure needs to be treated promptly, otherwise permanent disability or death may occur. The valves may function inappropriately due to calcification or other diseases and may need to be replaced. Clinical problems are also associated with blood vessels. These may become partially or completely blocked because of atherosclerosis. When this happens to the coronary arteries, bypass surgery with a grafted blood vessel may be required. Atherosclerosis, caused by high concentrations of cholesterol in the blood, leads to inappropriate changes in the plasma lipoproteins. A change of diet and drugs may help this condition. The blood pressure may be excessively high in some individuals. This can be dangerous and lead to heart failure and stroke. Here, again, drugs may be used to alleviate the condition. The peripheral blood vessels can also become diseased. For example, the veins returning blood to the heart may develop faulty valves, leading to varicose veins.

QUESTIONS

1. Which of the following is the odd ONE out?
(a) Endocardium.
(b) Bundle of His.
(c) Mitral valve.
(d) Pericardium.
(e) Semilunar valve.

2. Hypertension is a risk factor for which of the following?
(a) Asthma.
(b) Cerebral haemorrhage.
(c) Narrowing of the aorta.
(d) A high blood cholesterol concentration.
(e) Venous thrombosis.

3. Complete this sentence. The lesions found in atherosclerosis…
(a) are initiated as a response to damage to the venous endothelium.
(b) can develop at sites of intact endothelium.
(c) do not contain smooth muscle cells.
(d) do not contain macrophages.
(e) do not contain thrombi.

4. Arrange the following two lists into their most appropriate pairings:

Blood pressure	Tricuspid valve
Bundle of His	Pacemaker
Coronary arteries	Ventricular tachycardia
Pericardial diseases	Sphygmomanometer
Palpitations	Purkinje fibres
Right ventricle	t-PA
SA node	Lowered blood pressure
Thrombolysis	Myocardial infarction
Vasodilation	Familial hypercholesterolaemia
Xanthoma	Aspirin, ibuprofen

5. Account for the development of hypertrophic cardiomyopathy in some patients with pheochromocytoma. You may find it helpful to consult *Chapter 7*.

6. If a patient is found to have high VLDL concentrations in the blood, what change in diet could help to lower the concentration?

7. Suggest why chylomicrons are not taken up by LDL receptors.

8. List (a) three modifiable and (b) three nonmodifiable risk factors for heart disease.

FURTHER READING

Ciarambino T *et al*. (2021) Cardiomyopathies: an overview. *Int J Mol Sci* **22**:7722. doi: 10.3390/ijms22147722.

de Pádua Borges R *et al*. (2021) Choosing statins: a review to guide clinical practice. *Arch Endocrinol Metab* **64**:639–653.

Ding WY *et al*. (2022) Relationship between lipoproteins, thrombosis, and atrial fibrillation. *Cardiovasc Res* **118**:716–731.

Kazi DS *et al*. (2017) Statins for primary prevention of cardiovascular disease: review of evidence and recommendations for clinical practice. *Med Clin North Am* **101**:689–699.

Lee CC *et al*. (2020) High-sensitivity-cardiac troponin for accelerated diagnosis of acute myocardial infarction: a systematic review and meta-analysis. *Am J Emerg Med* **38**:1402–1407.

Libby P (2021) The changing landscape of atherosclerosis. *Nature* **592**:524–533.

Milman B and Burns BD (2021) Atrial fibrillation: an approach to diagnosis and management in the emergency department. *Emerg Med Pract* **23**:1–28.

Muntner P *et al*. (2019) Measurement of blood pressure in humans: a scientific statement from the American Heart Association. *Hypertension* **73**:e35–e66.

Rajani R and Klein JL (2020) Infective endocarditis: a contemporary update. *Clin Med (Lond)* **20**:31–35.

Schade DS *et al*. (2020) Cholesterol review: a metabolically important molecule. *Endocr Pract* **26**:1514–1523.

Shaya GE *et al*. (2022) Coronary heart disease risk: low-density lipoprotein and beyond. *Trends Cardiovasc Med* **32**:181–194.

Sinnenberg L and Givertz MM (2020) Acute heart failure. *Trends Cardiovasc Med* **30**:104–112.

Wang XY *et al*. (2020) The biomarkers for acute myocardial infarction and heart failure. *Bio Med Res Int* **2020**:2018035. doi: 10.1155/2020/2018035.

Useful website

www.abpi.org.uk

Genetic diseases

15

OBJECTIVES

After studying this chapter you should be able to:

- outline the major features of the cell cycle
- explain how genetic traits, including genetic diseases, are inherited
- describe the general effects of single gene mutations
- list the ways in which single gene disorders are treated
- list the types of chromosomal mutations/aberrations
- outline how genetic diseases are screened *in utero* and in neonates

15.1 INTRODUCTION

Genes are the fundamental units of heredity and encode specific functional products, such as RNA molecules and polypeptides. They are encoded by sequences of bases in DNA molecules and are found at particular positions in **chromosomes** in the nucleus and also in the relatively small circular DNA molecules in the mitochondria. Only 37 of the approximately 22 000 human genes occur in mitochondria, although mutations of these may become clinically significant as described in *Chapter 16*. Genes constitute the blueprint or the set of instructions that affects hereditary characteristics, for example hair and eye colour, height and the susceptibility to certain diseases.

When a cell divides, the genetic information needs to be replicated accurately so that these instructions pass on to the daughter cells. When changes occur in the base sequence of DNA, either as a result of incorrect replication or from random changes caused by physical or chemical agents, then the instructions become corrupted. This is a **mutation** (sometimes referred to as variants by clinicians) and may eventually lead to disease because the cell is unable to make, for example, a particular enzyme, signalling molecule, transporter or structural protein.

Chromosomes have a complex structure (*Figure 15.1*). Each comprises a double-stranded DNA molecule associated with numerous proteins. In general, chromosomes occur in matching or homologous pairs, with each member of a pair containing **alleles** or different forms of the same gene, which are found at identical loci (singular locus). Human somatic (body) cells contain 23 pairs of chromosomes and are said to be **diploid** (2N). The 46 chromosomes comprise 22 homologous pairs of **autosomes** (nonsex chromosomes) and one pair of sex chromosomes: XX in females and XY in males (*Figure 15.2*; see also *Figure 1.15*). Oocytes and spermatozoa have half the diploid number and are said to be **haploid** (N); oocytes can only contain an X chromosome, but a spermatozoan can have an X or a Y.

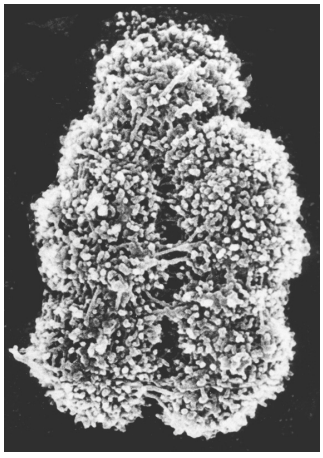


Figure 15.1 Scanning electron micrograph of a chromosome. Courtesy of Dr C.J. Harrison, Christie Hospital, Manchester, UK. ⏎

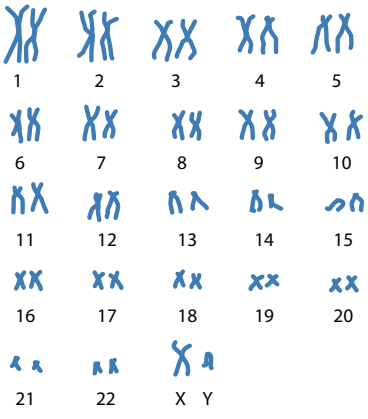


Figure 15.2 The chromosome complement of a normal human male. That of a normal female is shown in *Figure 15.23*. ⏎

DOI: 10.1201/9781003296836-15

15.2 GENETICS AND DNA

Genetics, from the Greek *genno* meaning 'to give birth', is the branch of biology concerned with heredity, genes and DNA (*Figure 15.3*), the genetic material. It is also the scientific study of the variations in inherited characteristics, often called traits, and how these are transmitted between generations. Inherited characteristics include a number of clinical conditions and diseases that are described in other chapters, for example inborn errors of metabolism (*Chapter 5*), sickle cell anaemia and haemophilia (*Chapter 13*), the muscular dystrophies and cystic fibrosis (*Chapter 16*), and some cancers (*Chapter 17*). Others, such as phenylketonuria (PKU) and Down syndrome, will be described in this chapter. **Genomics** is the study of the full complement of bases in the DNA of an organism.

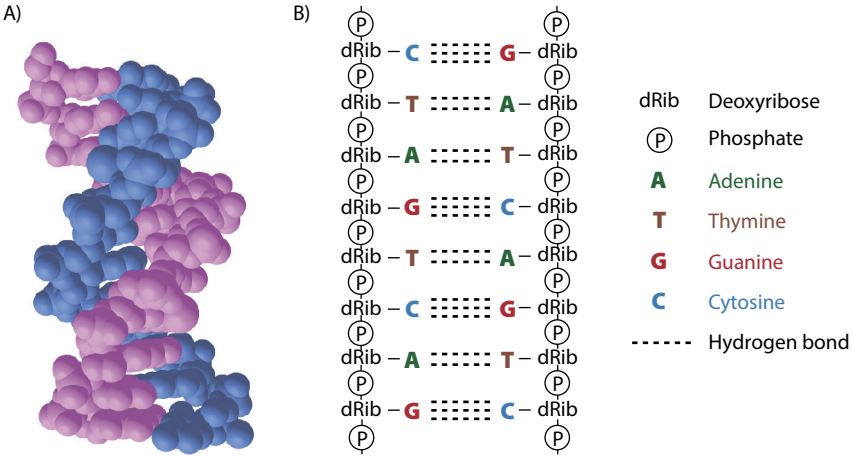

A)
B)
dRib – C ····· G – dRib
dRib – T ····· A – dRib
dRib – A ····· T – dRib
dRib – G ····· C – dRib
dRib – T ····· A – dRib
dRib – C ····· G – dRib
dRib – A ····· T – dRib
dRib – G ····· C – dRib
P
dRib Deoxyribose
Phosphate
A Adenine
T Thymine
G Guanine
C Cytosine
Hydrogen bond

Figure 15.3 (A) Molecular model of DNA, showing its two strands in contrasting colours. PDB file 4BNA. (B) Schematic illustration of the structure of DNA.

Genes are sequences of bases in DNA that carry the code for making RNAs or proteins. The code is contained in sequences of the four nucleotide bases: adenine, cytosine, guanine and thymine (A, C, G and T, respectively). DNA normally occurs as the famous double-helical molecule (*Figure 15.3*) that consists of two polymers of alternating phosphates and deoxyribose sugars, which are linked together by hydrogen bonds between pairs of complementary bases, rather like the steps of a spiral staircase. Adenine always pairs with T and C always pairs with G (*Figure 15.3*).

For a gene to be expressed, the nucleotide base code in a gene in the DNA must be copied or transcribed to form an RNA molecule. The strands of the double-helical DNA separate and one of them, 'the gene' or noncoding strand, is transcribed and forms the template for the biosynthesis of a complementary RNA strand (*Figure 15.4*). The base-pairing rules are similar to those in DNA, but uridine (U) is used not T; also, the sugar ribose, not deoxyribose, is found in RNA. Several RNA molecules may be transcribed from one gene. In some cases, the formation of an RNA molecule is the major event in gene expression. However, in the vast majority of cases, the expression of a gene results in the formation of a protein (*Figure 15.4*). In these cases, the RNA formed by transcription is a messenger RNA or mRNA. However, DNA is transcribed in the nucleus but the biosynthesis of proteins takes place in the cytoplasm. Thus, the mRNA molecules are

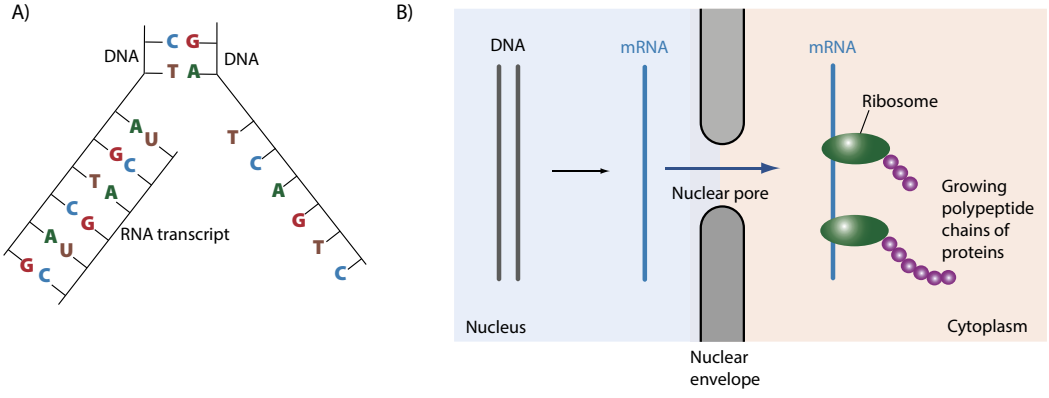


Figure 15.4 (A) A schematic illustrating the transcription of a DNA to form an RNA molecule. (B) An overview of translation (protein synthesis).

transported from the nucleus to the cytosol (*Chapter 16*). Here, their message is translated by ribosomes to give a linear sequence of amino acid residues forming a polypeptide, which is often modified and then folded to form a protein. Each sequence of three bases in the mRNA codes for the addition of one specific amino acid to the growing polypeptide chain of the protein; for example, AUG codes for the amino acid methionine, UUU codes for phenylalanine and so on. Each mRNA molecule may be translated numerous times so that many protein molecules are produced.

15.3 DNA REPLICATION AND THE CELL CYCLE

When a cell divides, the genetic information in DNA molecules must be passed on to the two daughter cells. The series of biochemical and morphological events that occur in a reproducing cell is called the **cell cycle**. This results in the replication of the DNA and division of the cell to give two daughter cells. Replication of DNA involves separating the two DNA strands of the double helix and aligning new bases according to the usual pairing rules: A with T and G with C. The new nucleotides are linked together to form two new strands, each of which is complementary to one of the original (parental) strands. The two new double-stranded DNA molecules each consists of one parental and one daughter strand (*Figure 15.5*) and, for this reason, is called semiconservative replication. This is a much simplified account of a complex process, which is catalysed by a range of enzymes and ancillary proteins. Replication of DNA is very accurate; it needs to be because the genetic instructions must be retained from generation to generation. Most of the few errors that inevitably occur are corrected by error-detecting enzyme systems in the cell.

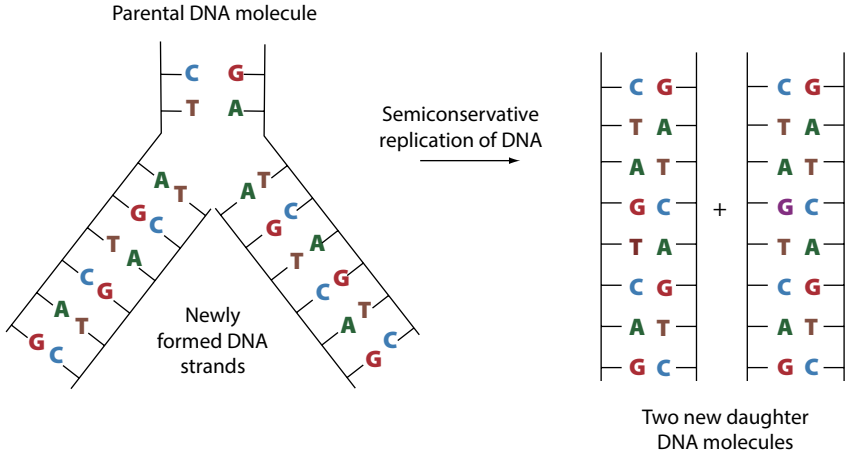


Figure 15.5 A schematic illustration of DNA replication, which emphasises its semiconservative nature.

Cell cycle

A **cell cycle** or cell division cycle is divided into two basic periods: **interphase** and **mitosis** (or **meiosis** in gamete production). The duration of a cell cycle varies in different types of eukaryotic cells but typically lasts about 24 hours. Approximately 95% of this time is spent in interphase, the period between mitoses, during which the cell prepares for division by building up large energy stores and synthesising new organelles for the daughter cells. A high metabolic rate is typical of cells about to undergo division.

The cell cycle can be divided into four main, but continuous, phases that are often drawn as a circle (*Figure 15.6*) called G_1, S, G_2 and M. The G_1 (for gap) phase lasts about eight hours, during which the cell makes a commitment to divide. It is characterised by the biosynthesis of RNA and protein. In the S (synthesis) phase, the DNA is replicated, which takes approximately six hours. The G_2 phase typically lasts four hours, during which organelles are replicated. Mitosis (or meiosis) occurs in the M phase. The M phase is followed by cytokinesis, when the cell divides to give two daughter cells.

Following mitosis, the cells enter interphase, which finishes when the S phase of the next cycle begins. Cells that divide only rarely, for example neurons, are said to be in a stage called the G_0 phase. It is only when they become committed to divide that they are described as being in the G_1 phase of the cycle.

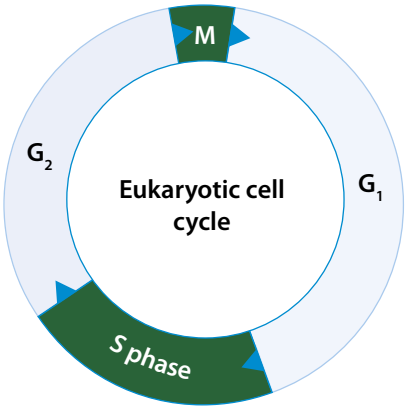


Figure 15.6 Overview of the stages of the cell cycle.

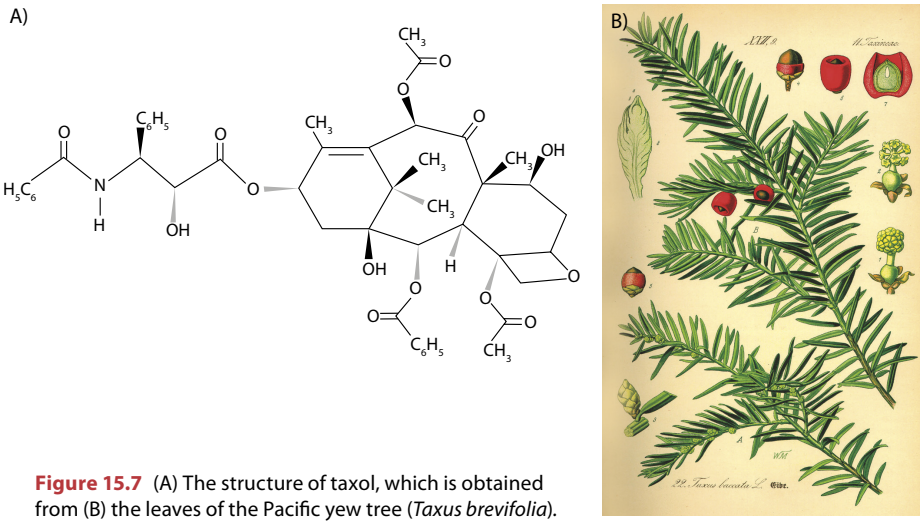


Figure 15.7 (A) The structure of taxol, which is obtained from (B) the leaves of the Pacific yew tree (*Taxus brevifolia*).

Mitosis and meiosis

Mitosis occurs during growth and the renewal of tissues, producing daughter cells with the same diploid complement of chromosomes as the parental cell. It is a continuous process (*Figure 15.8*) that lasts about one hour but for convenience it is divided into four stages:

1. prophase
2. metaphase
3. anaphase
4. telophase

In prophase, the chromosomes condense to form distinct mitotic figures that are visible using light microscopy. These contrast with the dispersed, thread-like chromosomal material called chromatin present during interphase and which cannot be seen by light microscopy. Since their DNA has been replicated, each chromosome is present as an identical pair of chromosomes, although at this stage each member of the pair is referred to as a sister **chromatid**, which are joined together by **centromeres** (*Section 15.7*). The centrosome, normally located just outside the nuclear envelope, is replicated and each one migrates to opposite poles of the cell. This allows the centrosomes to rearrange the microtubules of the cytoskeleton (*Chapter 16*) and form a mitotic spindle. The spindle spans the cell; its ends are known as poles, whereas its middle region is called an equator. During prophase, the cell's nucleolus disappears and it concludes with dissipation of the nuclear envelope. In metaphase, the chromosomes migrate to the centre of the cell and are arranged around the equator of the spindle, where the centromere of each chromosome (paired chromatids) becomes attached to spindle fibres. The chromatids are drawn apart at the centromere region towards opposite poles of the spindle. During anaphase, the chromatids of each chromosome are pulled apart and each chromatid moves towards the opposing poles of the spindles reaching the ends by telophase. In telophase, the cell undergoes division to two daughter cells by the plasma membrane constricting and cutting across the spindle equator. The spindle breaks down and nuclear envelopes form around each separated group of chromatids, now called chromosomes. The nucleoli also become apparent in the new nuclei and the chromosomes return to the nonvisible forms typical of interphase cells. Hence, the parental cell has divided to form two daughter cells that are genetically identical to the parent.

Meiosis occurs prior to reproduction during the formation of gametes (*Chapter 7*). Parental cells are diploid, whereas the daughter cells are now gametes with the corresponding haploid number of chromosomes. Given that the daughter cells receive only one of each type of chromosome, meiosis is sometimes called reduction division.

Meiosis may be thought of as consisting of two separate divisions (*Figure 15.9*). In the first meiotic division (prophase I to telophase I) the parent cell divides into two cells, each of which receives *one* of each pair of homologous chromosomes. Each of these chromosomes consists of two chromatids. The second meiotic division (prophase II to

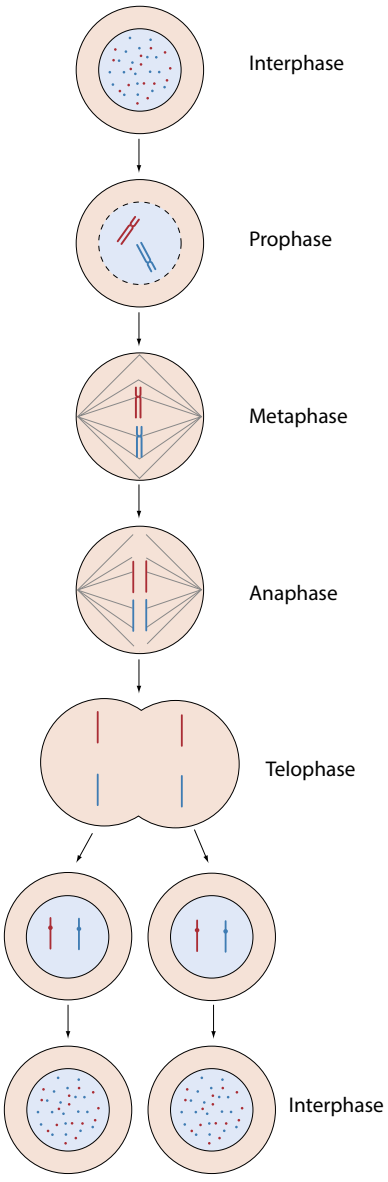


Figure 15.8 An outline of mitosis. For simplicity, only two chromosomes, each consisting of two chromatids, are shown.

Margin Note 15.1 The cell cycle and malignancy

The timed sequence of events in the cell cycle is a highly regulated process controlled by numerous different proteins. Mutations in the genes encoding these proteins can perturb the regulation of the cell cycle: cells may die or their growth may become uncontrolled depending on the nature of mutation. The latter may lead to malignancies (*Chapter 17*). Several chemotherapeutic drugs act by interfering with the cell cycle. For example, some immunosuppressant drugs, such as methotrexate and cyclophosphamide, disrupt the synthesis of DNA during the S phase. Taxol (*Figure 15.7*) is a drug that interferes with mitosis by preventing depolymerisation of the microtubules that make up the mitotic spindle.

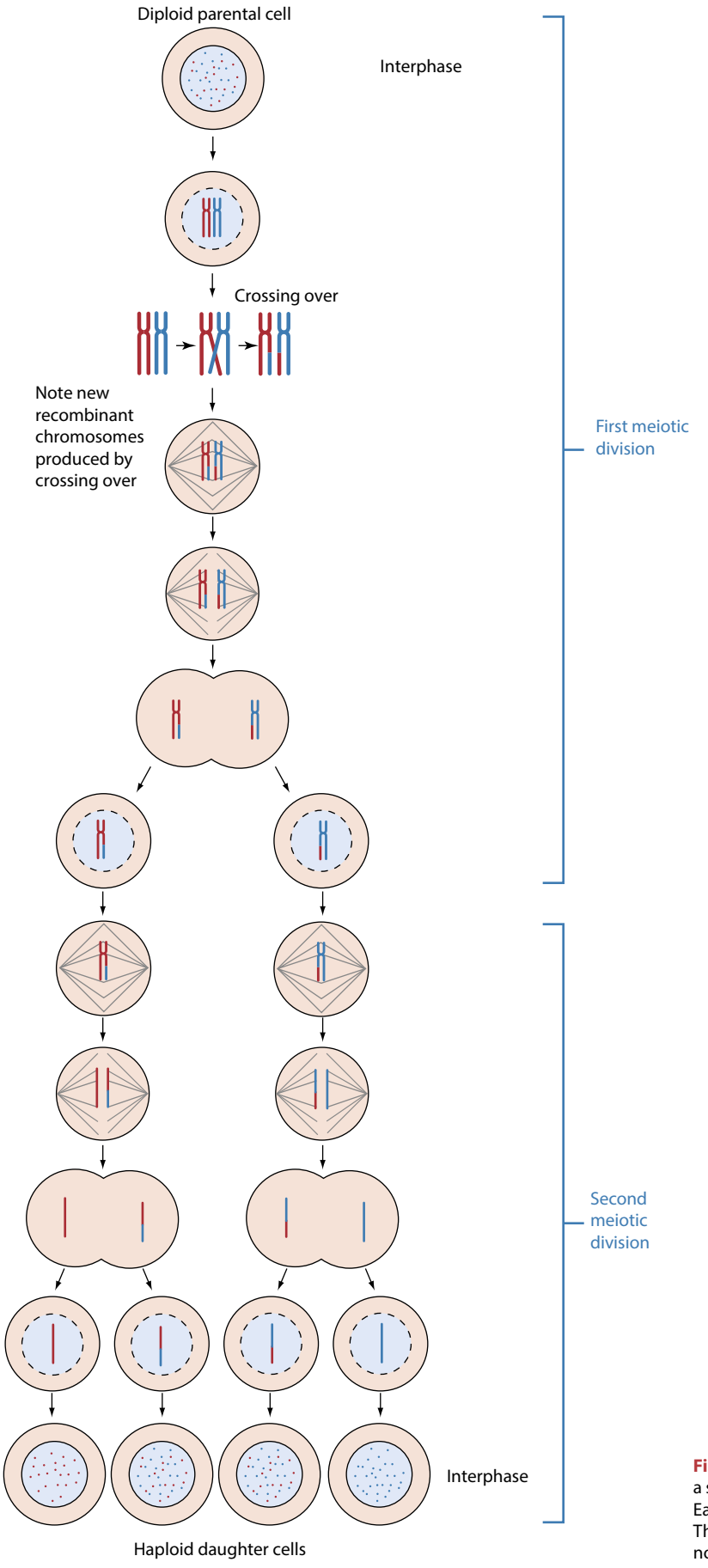


Figure 15.9 An outline of meiosis. For simplicity, only a single pair of homologous chromosomes is shown. Each member of the pair consists of two chromatids. The insert illustrates recombination between the two nonsister chromatids. ⏎

telophase II) results in each chromosome being separated into chromatids, giving four haploid daughter cells.

In prophase I, the chromosomes contract and the nucleolus shrinks. Homologous chromosomes lie side by side in pairs, a situation called synapsis. Each member of the pair is bivalent. Genetic recombination or crossing over occurs at this stage. Nonsister chromatids, that is, one maternal and one paternal chromatid of a homologous pair, are broken at equivalent positions and exchange homologous pieces of material (*Figures 15.9* and *15.10*). The crossed strands of the chromatids formed during recombination are called chiasmata (singular chiasma). Recombination results in each chromatid acquiring genes or parts of genes from the other, a process that results in the formation of new combinations of genes or parts of genes in chromosomes of the gametes. Thus, crossing over promotes genetic variation.

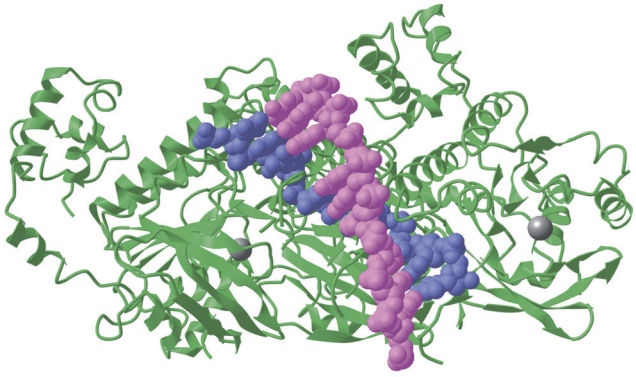


Figure 15.10 Molecular model of the human RAD51–DNA complex. RAD51 catalyses the ATP-dependent exchange of DNA strands during meiotic recombination. The protein subunits of the complex are shown in green, the DNA is in blue and magenta, and Mg^{2+} is shown as grey spheres. PDB file 5H1C. ⏎

During metaphase I, homologous chromosomes or bivalents move to the equator of the spindle; sister chromatids are orientated towards the same pole but the homologous chromosomes towards opposite poles. During anaphase I, the homologous chromosomes, each of course consisting of a pair of chromatids, migrate towards opposite poles of the cell. In telophase the cell divides, as in mitosis, to give to two daughter cells whose chromosomes each consist of paired chromatids. Following a brief interphase, these cells enter the second meiotic division. In prophase II, the two daughter cells essentially prepare for the second division with formation of a new spindle. During metaphase II, the chromosomes move to the equator of the spindle and the chromatids are oriented towards opposite poles and, in anaphase II, the chromatids separate from each other and move to opposite poles of the cell. Finally, in telophase II, each cell divides into two daughter cells but these have only the haploid number of chromosomes. Thus, the diploid parental cell has produced four haploid daughter cells (spermatozoa or ova [*Chapter 7*]).

Nondisjunction is the failure of chromosomes to separate during metaphase in mitosis or either of the metaphases in meiosis and can have serious clinical consequences as explained in *Section 15.8*.

15.4 GENOTYPE AND PHENOTYPE

The complete genetic or hereditary constitution of an individual forms his/her **genotype**. This term can also be applied to any particular pair of corresponding alleles in a chromosome. In contrast, the visible or measurable characteristics of an individual constitute the **phenotype**. A phenotype includes biochemical, physiological, morphological and behavioural characteristics or, indeed, any observable biological trait, such as the total physical appearance and constitution of an individual or any specific trait, such as size, weight or eye colour and, of course, includes characteristics of clinical importance and the presence of a clinical condition. Some phenotypic traits, for example eye colour, are directly observable but others, such as the blood group of a patient (*Chapter 6*), may only become apparent following specific tests. Phenotypic traits do not necessarily occur merely following the expression of the genotype of an individual; some, such as the blood groups, are completely determined by heredity, but many others, for example weight and height (*Chapter 10*), result from interactions between the genotype and the environment. Phenotype can also alter due to development changes, for example when embryonic, then fetal, followed by adult haemoglobin is progressively expressed (*Chapter 13*).

15.5 INHERITANCE AND MUTATIONS

Each allele of a gene is carried by two respective homologous chromosomes. If the two alleles are identical, the individual is **homozygous** for that gene; if they differ, the individual is **heterozygous**. In the heterozygous state, only one allele may be expressed and is said to be **dominant** over the other, which is **recessive**. The recessive trait will only become apparent in a homozygous recessive individual.

Dominant genes are conventionally written as an upper case italic letter, for example *G*, while its recessive counterpart is given in the lower case form, *g*. *Figure 15.11* illustrates the normal inheritance pattern first established by Gregor Mendel (1822–1884). If one parent is homozygous for an autosomal dominant gene (*GG*) and the other is homozygous for the recessive form (*gg*), then all the offspring will be genetically heterozygous (*Gg*) and phenotypically will express the dominant trait. If both parents are heterozygotes (*Figure 15.11*), then statistically 25% of the offspring will be homozygous for the dominant gene (*GG*), 25% will be homozygous for the recessive gene (*gg*) and the remaining 50% will be heterozygous (*Gg*).

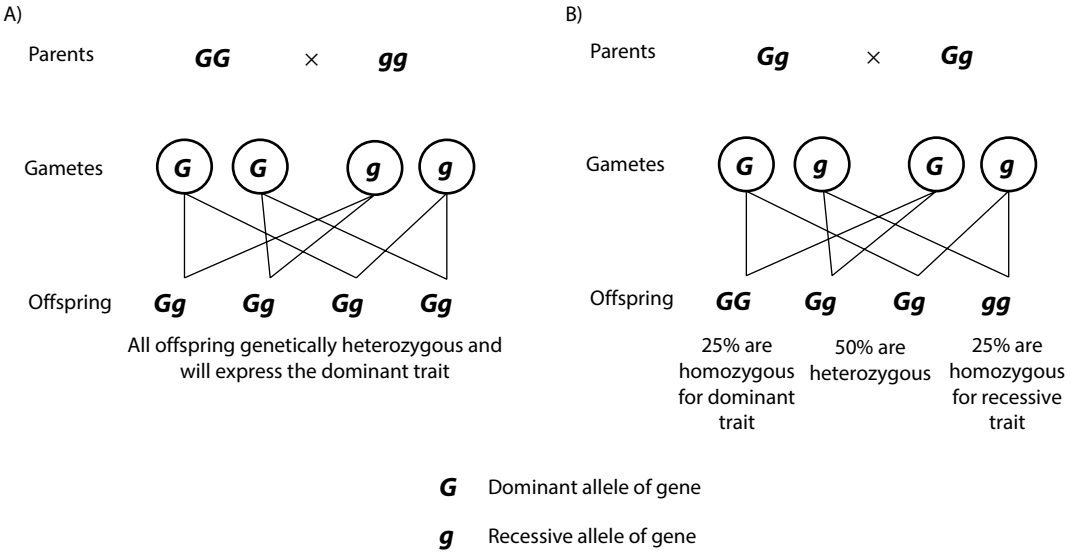


Figure 15.11 The inheritance patterns shown for a single pair of genes between (A) two contrasting homozygous parents and (B) two heterozygous parents.

Mutations are changes in the **genome** (*Margin Note 15.2*) and can give rise to clinical disorders. Mutations vary in character. They may simply be the substitution of one nucleotide for another, called point mutations, or involve the insertion or deletion of one or more nucleotides within the normal sequence of DNA or alter the structures of individual chromosomes or even the number of chromosomes present.

When considering the effects of mutations, it is necessary to distinguish between those that occur in somatic cells and those affecting gametes. In somatic cells, mutations that produce a recessive autosomal allele are unlikely to have clinical consequences because their expression is masked by the corresponding dominant allele. However, somatic mutations that are dominant or X-linked (see below) can have a greater impact because they are likely to be expressed. Similarly, their impact is greater if they arise early in development before undifferentiated cells give rise to differentiated tissues or organs. In adult tissues, the activities of many nonmutant cells often mask mutations in a few other cells. Mutations in somatic cells will not be transmitted to future generations although they may represent the first step in the development of cancer (*Chapter 17*).

Mutations in gametes or gamete-forming tissues are part of the germ line (*Chapter 17*) and are of greater clinical concern because, as well as affecting that individual, they can also be transmitted to offspring.

A dominant autosomal mutation (*M*) will be expressed phenotypically in both the homozygous and the heterozygous conditions. However, if the mutation is recessive (*m*), then it is not likely to affect an individual unless both chromosomes carry the same mutation (*Figure 15.12*). The inheritance patterns shown by dominant/recessive alleles that are associated with a clinical condition follow Mendelian rules (see above). If only a single member of a homologous pair of chromosomes carries the mutation, it can, however, be passed to the next generation and the parent is described as a carrier. If one parent is homozygous for the normal gene (*MM*) and the other heterozygous for the

Margin Note 15.2
The human genome

The **genome** is the complete sequence of bases in the DNA molecules, which is all of the hereditary material possessed by an individual. The human haploid genome contains about 3 000 000 000 pairs of nucleotides, whose sequence was reported in 2001. The total length of this DNA is about a metre and is divided into 23 individual molecules, 22 of which are found in the autosomes and one in each of the sex chromosomes, which together with mitochondrial DNA (*Chapter 16*) contains approximately 22 000 genes.

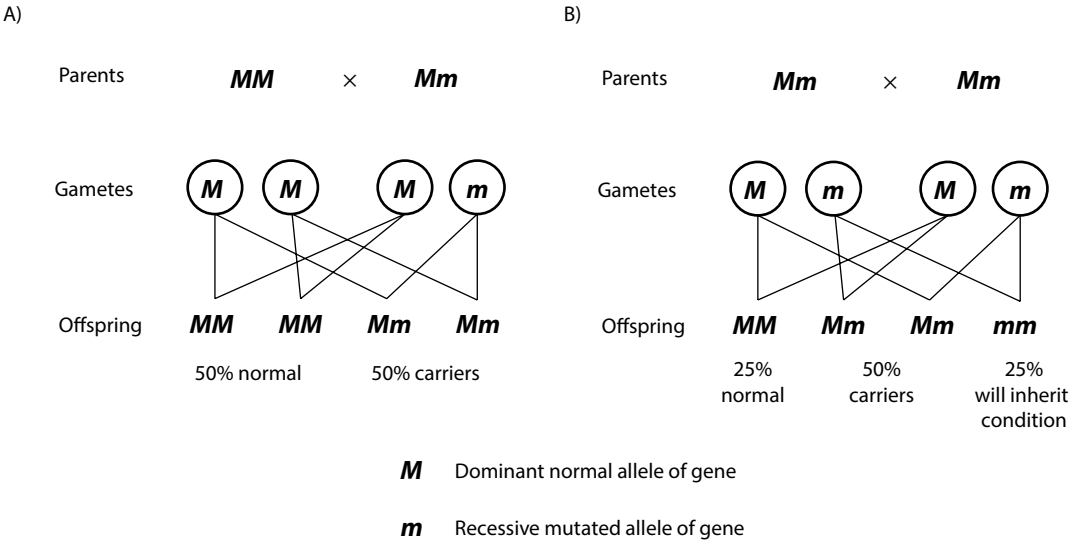


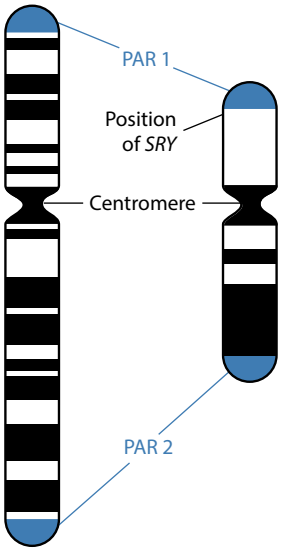
Figure 15.12 The inheritance patterns shown by a recessive allele of a gene between (A) dominant homozygous and heterozygous parents and (B) two heterozygous parents.

mutated allele (*Mm*) then 50% of the offspring will be homozygous for the normal gene (*MM*) and the other 50% will be heterozygous (*Mm*) and are carriers. None of the children will be affected. If both parents are heterozygotes (*Mm*), then 25% of the offspring will be dominant homozygotes (*MM*) and 50% will be heterozygotes and carriers (*Mm*), but the remaining 25% will be recessive homozygotes (*mm*) and express the condition.

Heterozygosity means that autosomal recessive mutations, even one resulting in a lethal allele, may go unnoticed and be maintained until the resultant allele is widespread in the population. The new allele will become evident only when a chance mating brings two copies of it together in the homozygous condition.

Sex-linked genetic diseases

A number of genetic diseases are caused by defective alleles of genes of the sex (X and Y) chromosomes (see, for example, *Chapter 5*). The X chromosome contains appreciably more genes than the Y (*Figure 15.13*), although they do have some genes in common; thus, any mutated gene on the X chromosome is likely to be expressed in males (XY) but be masked in females (XX). Genetic diseases associated with the X chromosome are



Size of DNA (bp)	~156 × 10^6	~57 × 10^6
Number of genes	~900	~55

Figure 15.13 Cartoons of the X and Y chromosomes. The X chromosome is about three times larger than the Y, which contains far fewer genes. PAR, pseudoautosomal region; *SRY*, male sex-determining gene on the Y chromosome.

most common and are often referred to as X- or sex-linked genetic conditions. X-linked diseases can be recessive or dominant, although the former, for example haemophilia (described in *Chapter 13*), are the better known (*Table 15.1*). The expression of sex-linked genes differs between females and males because many genes on the Y chromosome lack a corresponding allele on the X chromosome. Hence, X-linked recessive genes are only expressed in females if there are two copies of the gene: one on each of the X chromosomes. For males, only one copy of a defective X-linked recessive gene is required for the disorder to be expressed. For example, if a woman carried a mutated recessive gene on one of the X chromosomes (*Figure 15.14*) then, statistically, 50% of her sons would inherit the defective gene.

Recessive	Dominant
Duchenne and Becker muscular dystrophies (*Chapter 16*)	Coffin–Lowry syndrome
Haemophilia (*Chapter 13*)	Incontinentia pigmenti
Red–green colour blindness	
Wiskott–Aldrich syndrome (*Chapter 5*)	
X-linked agammaglobulinaemia (*Chapter 5*)	

Table 15.1 Examples of X-linked diseases.

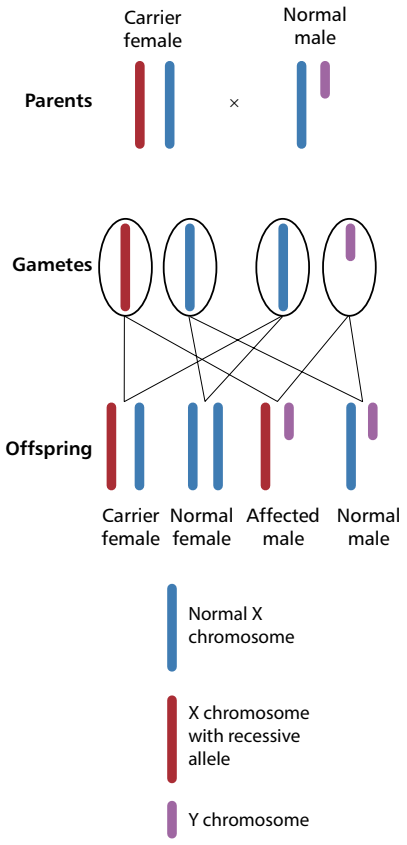


Figure 15.14 The inheritance pattern shown by a recessive allele carried on one of the X chromosomes of the mother with a normal male father.

15.6 INHERITED GENE DISORDERS

Most inherited diseases are due to mutations in genes in the nuclear chromosomes, although they can also occur as a result of mitochondrial gene mutations (*Chapter 16*). The mutation of a single gene may lead to the absence or modification of a specific protein, for example the abnormal haemoglobin in sickle cell anaemia (*Chapter 13*). In some cases, the inherited disorder may result in defective biosynthesis of a receptor, such as in familial hypercholesterolaemia (*Chapter 14*) where there is a defect in low-density lipoprotein (LDL) receptors, or in dysfunctional carrier proteins, such as in cystinuria where renal reabsorption of cystine (formed by the oxidation of two cysteines) is impaired. If the defective or absent protein is an enzyme, the result is a metabolic disorder. Most inherited metabolic disorders are autosomal recessive diseases with symptoms seen only in the homozygous condition, although heterozygotes are phenotypically normal because sufficient amounts of the protein are produced. Nevertheless, a number of these conditions have an autosomal dominant mode of inheritance and consequently heterozygotes are affected. Examples of these include porphyrias (*Chapter 13*) and familial hypercholesterolaemia (*Chapter 14*).

Consequences of an enzyme deficiency

In inherited metabolic disorders caused by a complete or partial deficiency of an enzyme that controls a particular reaction in a metabolic pathway, the clinical features are a consequence of the enzyme deficiency. *Figure 15.15* shows how this may occur. Symptoms arise from a lack of product D if this is an essential substance, and alternative pathways for its biosynthesis do not exist. Moreover, an accumulation of precursor C will occur if enzyme activity is absent and may produce clinical features if the accumulated substrate is toxic when it accumulates. Intermediates of the pathway may also build up, especially since there will be no negative feedback effect, as the final product of the pathway is absent. Finally, accumulation of a product of a minor pathway, E, may occur and, if this is toxic in excess, may also produce clinical features.

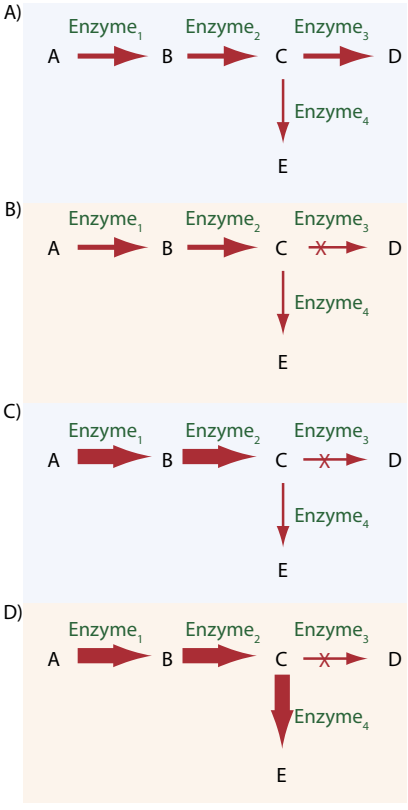


Figure 15.15 Idealised metabolic pathways to show effects of an enzyme deficiency. Metabolites are shown in upper case letters, while the relative rates of the reactions are indicated by the size of the arrows. (A) The metabolic pathway illustrates synthesis of product D from substrate A by a series of reactions catalysed by enzymes 1, 2 and 3. Product E is derived from a minor pathway by the action of enzyme 4 on substrate C. A deficiency of enzyme 3 has a number of consequences including (B), where conversion of C to D is blocked resulting in a decline in the production of D. In (C), the concentration of B and C increases due to the increased activities of enzymes 1 and 2, because the lack of D means there is no negative feedback on enzyme 1 and possibly 2, and (D) shows the increase in the conversion of C to E by enzyme 4.

The treatment of inherited metabolic disorders is to try and prevent any accumulation of precursor(s) and to provide the necessary product of the pathway. The removal of toxic products of any minor pathways may also be necessary. Future strategies are aimed at replacing the deficient enzyme or correcting the defective gene by gene editing.

Many inherited metabolic disorders caused by a deficiency of an enzyme are known (*Table 15.2*); one of the most thoroughly documented is PKU.

Disorder	Enzyme deficiency
Phenylketonuria (PKU; see below)	Phenylalanine hydroxylase
Porphyria cutanea tarda	Uroporphyrinogen decarboxylase
Gaucher disease (*Chapter 16*)	Glucocerebrosidase
Forbes disease	α,1,6-Glucosidase
Fabry disease	α Galactosidase
Glycogen storage diseases	Deficiency of one of several enzymes involved in interconverting glycogen and glucose
G6PD deficiency (*Chapter 13*)	Glucose-6-phosphate dehydrogenase (G6PD)

Table 15.2 Examples of inherited enzyme deficiencies. ⏎

BOX 15.1 SPONGIFORM ENCEPHALOPATHIES OR PRION DISEASES

Spongiform encephalopathies (SEs) or prion diseases (*Table 15.3*) are a rather peculiar group of diseases that can be inherited, although this is not their usual method of transmission. It was noted in *Chapter 2* that these are infectious diseases but also have a low sporadic occurrence of about one in a million for Creutzfeldt–Jakob disease (CJD). They are unusual in that they can also be familial, that is, they are also inherited (genetic) diseases. Since the diseases usually occur only after reproductive life is over, they can run in families, giving the familial form. The incidence of inherited cases appears fairly static and, although sporadic cases appear to be increasing, this may simply be a facet of more accurate reporting (*Figure 15.16*). Irrespective of cause, these diseases generally develop slowly over 10–20 years in older individuals and are characterised by the presence of holes or plaques in brain tissue, which can be observed by some scanning techniques or postmortem, giving it a spongy appearance (*Figure 15.17*), hence the name spongiform. Cures are not available for these diseases and all are fatal. Variant CJD, which first appeared in the UK in the 1980s, differs from 'conventional' CJD in that it occurs in younger people and death occurs relatively rapidly, within about two years, following the first appearance of symptoms (*Chapter 2*).

Spongiform encephalopathies
Atypical dementias
Variant Creutzfeldt–Jakob disease (vCJD)
Fatal familial insomnia (FFI)
Gerstmann–Sträussler–Scheinker disease (GSS)
Kuru

Table 15.3 Examples of human spongiform encephalopathies or prion diseases. ⏎

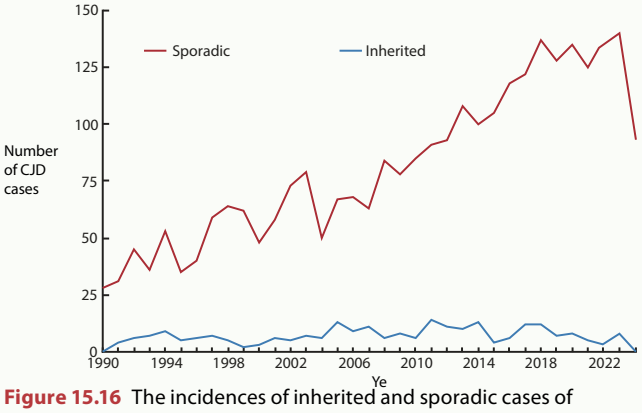


Figure 15.16 The incidences of inherited and sporadic cases of Creutzfeldt–Jakob disease (CJD) during 1990–2024. Data from https://www.cjd.ed.ac.uk. ⏎

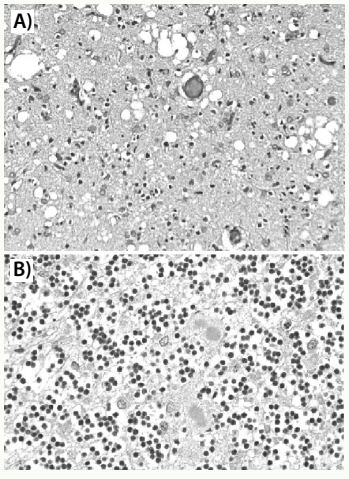


Figure 15.17 (A) The distinctive spongiform appearance of the cortex of the brain associated with Creutzfeldt–Jakob disease (CJD) and (B) aggregates of the pathological form of the protein deposited in the cerebellum. Courtesy of the National CJD Surveillance Unit, UK. ⏎

Prion proteins occur in a predominantly α helical conformation, the native form. However, these molecules can change shape to a form with an increased β sheet content that is a pathological conformation (*Figure 15.18*). The β sheet-rich prion protein induces conformational changes in native α helical-rich molecules, changing them to the β-rich form. These newly misfolded molecules can, in turn, stimulate conformational changes in other molecules in a chain reaction that deposits aggregates of prions in the brain leading to the destruction of neurons and finally the lethal spongiform condition. Thus, SEs are a subdivision of a group of diseases called *protein conformational diseases*.

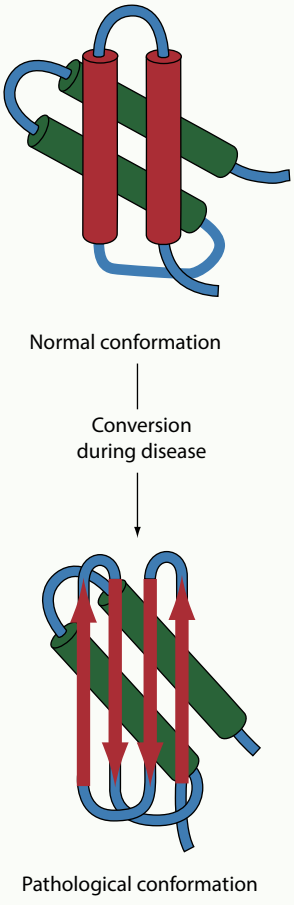


Figure 15.18 Schematic showing the change from the normal (α helical-rich) prion protein to the pathological (β sheet-rich) conformation.

The sporadic forms of these diseases occur in individuals with point mutations in the gene that encodes the prion protein. These mutations alter the sequence of amino acid residues in the prion protein molecules and predispose them to misfold to the β sheet-rich, pathological form. Different mutations in the gene are associated with different SEs (*Table 15.4*). The mutations are, of course, heritable, but since the symptoms usually only become apparent after reproductive life is over, they can run in families, producing the familial forms of the disease.

Mutation	Disease
Pro102Leu	GSS
Pro105Leu	GSS
Ala117Val	GSS
Tyr145Stop	GSS
Asp178Asn	Familial CJD, FFI
Val180Ile	GSS
Phe198Ser	GSS
Glu200Lys	Familial CJD
Arg208His	CJD
Val210Ile	Familial CJD
Gln217Arg	GSS
Octarepeat insert	Familial CJD

CJD, Creutzfeldt–Jakob disease; FFI, fatal familial insomnia; GSS, Gerstmann–Sträussler–Scheinker disease.

Table 15.4 Some mutations of the human prion protein associated with spongiform encephalopathies.

Recent years have seen the development of effective methods of diagnosing these conditions, of which magnetic resonance imaging (*Figure 15.19*) and tests for protein markers, such as tau in cerebrospinal fluid, are noteworthy. Sadly, even when accurately diagnosed, there are no effective treatments and these conditions are invariably fatal.

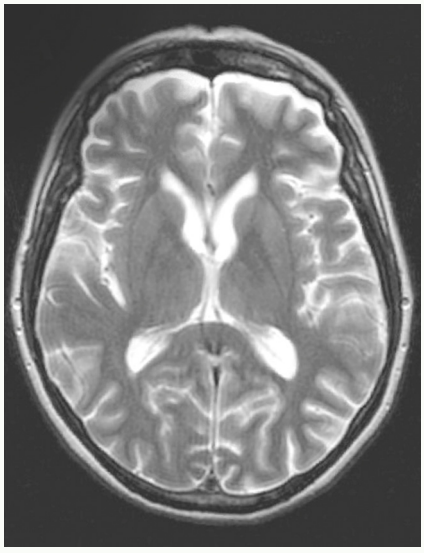


Figure 15.19 Magnetic resonance image of the brain of a patient with sporadic Creutzfeldt–Jakob disease (CJD). Courtesy of Dr P. Rudge, The National Prion Clinic, UK. See also *Figure 2.5*.

Phenylketonuria

Phenylketonuria is the most common disorder of amino acid metabolism, with an incidence of 1 in 10 000 in the UK; in other countries, it varies from 1 in 5000 to 20 000 births. It is an autosomal recessive condition that results in a deficiency, mostly in the liver, of phenylalanine hydroxylase, which catalyses the hydroxylation of phenylalanine to tyrosine. Tyrosine is required for the synthesis of proteins, the pigment melanin, thyroxine and the catecholamine hormones (*Chapter 7*). However, if phenylalanine hydroxylase activity is absent or greatly diminished, then phenylalanine and its metabolites accumulate and are toxic to the developing brain. There is also an increase in the level of phenylpyruvate, a phenylketone, which is normally a minor metabolite of phenylalanine (*Figure 15.20*). Excess phenylpyruvate is excreted in urine, hence the name phenylketonuria.

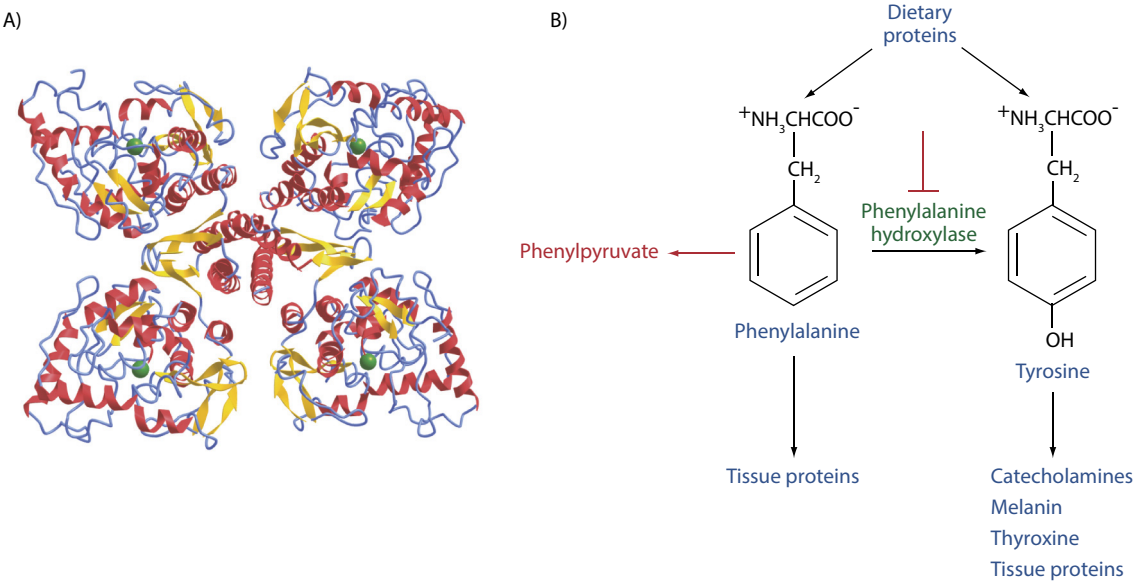


Figure 15.20 (A) Molecular model of phenylalanine hydroxylase. The green spheres represent Fe^{2+}. PDB file 2PAH. (B) The metabolism of phenylalanine. The reaction inhibited by a diminished level or absence of phenylalanine hydroxylase in phenylketonuria is indicated. ⏎

The clinical features of PKU are absent at birth but develop within a few days if the newborn is untreated. These signs include a characteristic mousy odour, irritability, poor feeding, vomiting, eczema and intellectual disabilities, as well as pale skin, fair hair and blue eyes due to decreased melanin biosynthesis. The most serious of these features is irreversible intellectual disability, which develops within three to six months following birth. A diagnosis of PKU is made when the concentration of phenylalanine in the serum is more than 0.7 mmol/L, compared with the reference range value of less than 0.1 mmol/L. Clinical determinations are performed a few days following birth, since it is vital to begin treatment as soon as possible.

The management of PKU involves restricting the dietary intake of phenylalanine so that the serum phenylalanine concentration does not exceed the limits shown in *Table 15.5*. In the early stages of life, when the brain is developing rapidly, strict control of phenylalanine concentrations must be imposed to prevent brain damage. This includes not feeding the baby breast milk, since this contains appreciable amounts of phenylalanine. Early treatment is crucial because the IQ of an affected individual rarely exceeds 70 and the brain damage caused by untreated PKU is irreversible.

Age (years)	[Phenylalanine] (mmol/L)
0–5	<0.36
5–10	<0.48
10+	<0.70

Table 15.5 Recommended limits of serum phenylalanine in phenylketonuria patients. ⏎

Commercially prepared diets with low amounts of phenylalanine are available, but the contents cannot be zero, since phenylalanine is an essential amino acid (*Chapter 10*), and so some must be provided in the diet to support protein biosynthesis. Tyrosine is not an essential amino acid unless phenylalanine intake is limited. Therefore, adequate quantities of tyrosine must also be provided to PKU patients. Children who have been diagnosed shortly after birth and properly treated by dietary management develop normally. Patients should be tested regularly to check that they are adhering to their diets. Regular monitoring of treatment is advisable. The diet can become somewhat less rigorous after the age of 10, although many clinicians now believe that dietary restriction should be continued throughout life. If dietary restriction is started early, the prognosis is very good, and many patients are approaching old age. Sapropterin is licensed for the treatment of children of all ages with PKU. This is a synthetic form of the phenylalanine hydroxylase cofactor tetrahydrobiopterin and is thought to enhance residual enzyme activity.

Neonatal screening programmes (*Margin Note 15.3*) for PKU are well established in practically all developed countries, including the UK. Any screening programme must be both cost effective and ethical. Screening for PKU involves collecting a specimen of capillary blood 6–10 days following birth, which allows for sufficient time for feeding and protein intake to become established, and determining the concentration of phenylalanine in the plasma. If the result is indicative of PKU, further definitive tests are performed. The plasma phenylalanine concentration used to be determined by the Guthrie test, which involves investigating the ability of the plasma to support the growth of *Bacillus subtilis*, which can only grow if phenylalanine is present in the medium; this was superseded by chromatographic, fluorometric methods. Nowadays, most laboratories use mass spectrometric methods to determine phenylalanine concentrations.

Strict dietary control is also necessary in pregnant women who have PKU, since maternal hyperphenylalaninaemia can affect the fetus *in utero*, even if the fetus itself does not suffer PKU. Intellectual disability and congenital abnormalities can occur in a large proportion of these infants.

Margin Note 15.3
National screening programmes for inherited diseases

Several factors need to be considered before a screening programme (*Chapter 1*) for any inherited disease is instituted. These include the following.

- Does the disease have a relatively high incidence?
- Can the disease be detected within days of birth?
- Can the disease be identified by a biochemical marker that is easily measured?
- Will a failure in diagnosing the disease early cause irreversible damage to the baby?
- Can the disease be treated, and will the result of any screening test be available before irreversible damage to the baby occurs?

15.7 CHROMOSOMES AND THE HUMAN KARYOTYPE

Interphase chromosomes are present as extended structures and cannot be seen by light microscopy. At the onset of both mitotic and meiotic cell division, the chromosomes condense to form compact structures, referred to as mitotic figures. All chromosomes have a narrowed region called the centromere that divides the chromosome into two portions and allows them to be classified as metacentric, submetacentric, acrocentric or telocentric as shown in *Figure 15.21*.

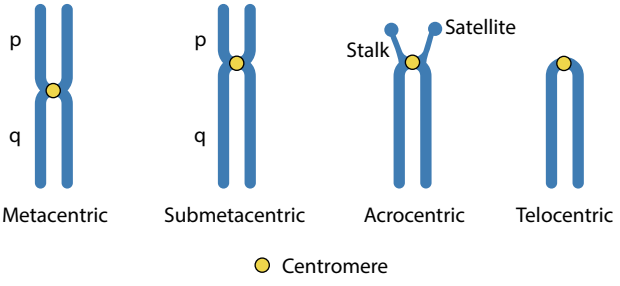


Figure 15.21 A classification of chromosomes into metacentric, submetacentric, acrocentric and telocentric types based on the position of the centromere.

When a chromosome is divided into unequal lengths, the shorter portion is called the *p* arm and the longer the *q* arm. When stained with Giemsa stain, both arms are divided into two or more regions by prominent bands and each region is further subdivided into subbands that can be numbered unambiguously (*Figure 15.22*). For example, band Xp21.2 is to be found on the *p* arm of the X chromosome in region 2, band 1, subband 2.

A **karyotype** is the characteristic number, size and shape of chromosomes of a species, while a **karyogram** is a photographic representation of these chromosomes stained and arranged in order (*Figures 15.2* and *15.23*). An **idiogram** is a diagrammatic representation or interpretive drawing of the chromosomes based on the physical features seen in the karyogram. The karyotype of normal humans is 46. Human autosomal chromosomes are divided into seven groups (*Table 15.6*) based on their sizes and the positions of their centromeres.

Figure 15.22 A single human A2 (see *Table 15.6*) chromosome showing G banding. Courtesy of J.S. Haslam and K.P. O'Craft, Tameside General Hospital, Ashton under Lyne, UK.

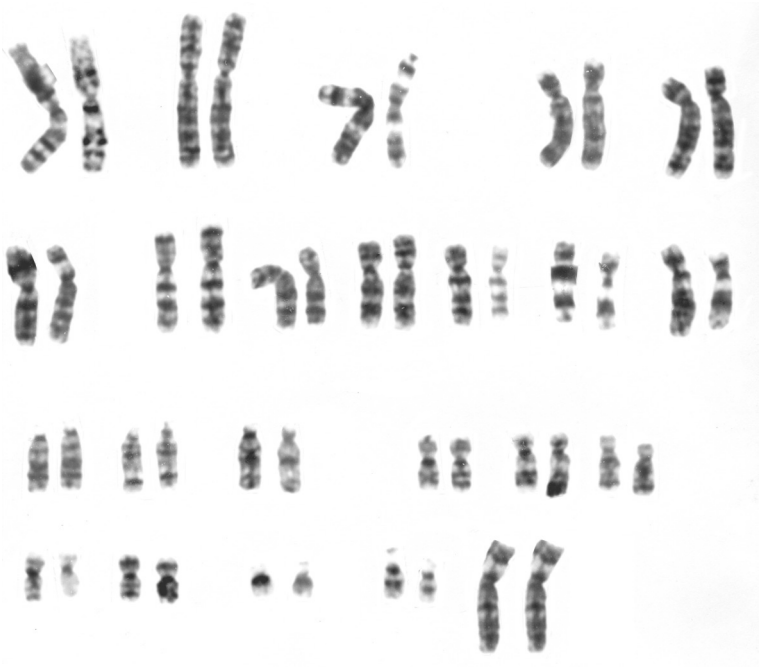


Figure 15.23 Karyogram of a normal human female showing G banding. Courtesy of J.S. Haslam and K.P. O'Craft, Tameside General Hospital, Ashton under Lyne, UK. ⏎

Chromosome group	Chromosome numbers	Structures
A	1–3	
B	4–5	
C	6–12, X	
D	13–15	
E	16–18	
F	19–20	
G	21–22, Y	

Table 15.6 The seven groups of human chromosomes. ⏎

Cytogenetics is the microscopic study of chromosomes. Small lymphocytes isolated from a blood sample or cells obtained by amniocentesis or chorionic villus sampling (*Section 15.9*) are stimulated to divide by treatment with the plant lectin phytohaemagglutinin (*Table 5.3*), and mitosis is then arrested at metaphase by a mitotic inhibitor, such as colchicine (see *Figure 17.34*). Metaphase chromosomes are visible by microscopy when stained in one of several different ways to allow their accurate identification when examined. For routine karyotyping, Giemsa (G) staining is usually the preferred procedure, since this produces a pattern of alternating dark and light bands that are characteristic for each pair of chromosomes, which reflects the differences in the detailed structure between each chromosome. Karyotyping can be performed on leukocytes from whole blood as described above or from amniotic fluid, which contains cells from the developing fetus. When patient chromosomes are examined by microscopy, it is possible to identify aberrations in chromosome number and structure.

The phenotype of a patient with a chromosomal disorder depends on the nature of their chromosomal defect. The first human chromosomal disorder was discovered in 1959, when three copies of chromosome 21 were found to be associated with Down syndrome (*Section 15.8*). The development of chromosomal banding since 1970 has markedly

increased the ability to resolve small chromosomal aberrations. Contemporary methods in molecular biology, for example next-generation sequencing of DNA (*Chapter 17*), means the specific mutation can rapidly be identified.

Fragile sites

When human cells are grown in culture (*Chapter 17*), some of the chromosomes in cells derived from certain individuals fail to stain in particular regions, giving the appearance of a gap known as a **fragile site** (*Figure 15.24*). These sites are susceptible to breakage during culture in the absence of certain chemicals such as folic acid, which is normally present in the culture medium. More than 120 such sites have been identified since they were first discovered in 1965. Almost all studies on fragile sites have been carried out *in vitro* on cells halted in mitosis. However, the cause of the fragility is not known with certainty, but they correlate with variations in copy number and chromosome rearrangements involved in tumour development and some neurological disorders. Human miRNAs (microRNAs; *Margin Note 15.4*) are encoded by DNA in fragile sites. The aberrant expression of miRNAs is associated with many human diseases. A strong association has been shown to exist between a form of intellectual disability called fragile X or Martin–Bell syndrome and a fragile site on the X chromosome at Xq27.3 associated with *FMR1* (fragile X mental retardation 1), a gene which codes for a protein involved in ribosomal translation. It is a dominant trait but, fortunately, fails to be fully expressed (referred to as 'incomplete penetrance') in many individuals. However, it is the most common inherited cause of reduced mental capabilities and has been estimated to affect 1 in about 4000 males in the UK. Most who suffer intellectual disability do so to the point that they are unable to live an independent life and have a distinct physical appearance, including long, narrow faces with protruding chins, enlarged ears and increased testicular size particularly after puberty. The syndrome also affects about 1 in 6000 females, who tend to suffer milder intellectual disabilities.

The patterns of inheritance of fragile X syndrome are somewhat complicated compared with a disorder like PKU (*Section 15.6*). Most humans carry a stable version of *FMR1* that has about 30 CGG repeats. Individuals who have *FMR1* with about 45–55 CGG repeats are in a grey zone; they do not have fragile X syndrome and, while they are likely to pass on a stable gene to their children, they have an increased chance of having children with a larger number of CGG repeats. People with approximately 55–200 are said to have a *premutation*, since, although they generally have few or lack symptoms of fragile X syndrome, they can have children with more than 200 CGG repeats. This is the full mutation that initiates inappropriate methylation of cytosine bases (*Box 15.2*) and, sadly, can lead to the full blown syndrome. The gene *FMR1* shows Mendelian patterns of inheritance for X-linked disorders (*Figure 15.14*), although if it reaches the premutation stage it has a high probability of mutating (by repeat amplification) from one generation to the next. Amplification of the CGG repeats only occurs in females, not males; thus, although there may be no family history of fragile X syndrome, it can suddenly appear in a number of offspring.

Given that *FMR1* is on the X chromosome, fathers cannot pass on any form of it to male offspring. However, daughters generally receive the paternal type. For example, almost all males with the stable version generally have daughters with the stable version. However, males with premutations, who are generally phenotypically normal and referred to as normal transmitting males, have premutation-type daughters. This inheritance can have severe consequences for any male grandchild as explained below. Most full mutation males do not have children. Those few who do give the full mutated version to their daughters but, surprisingly, the daughters only express the premutation. Hence, the father is passing on a reduced number of CGG repeats, presumably because there are protected cells in the testes that never expand to the full mutation or a reduction in the repeat number occurs in some male reproductive cells. This means that all females who do have the full mutation must have received it from their mothers, since they cannot receive it from their fathers.

Given that females have two X chromosomes, every child, male or female, has an equal, random chance of receiving one or the other of them. A female who has a copy of the premutation from her normal transmitting father can pass it on to her children. Most daughters who receive the premutation will show an increase in repeat number compared with their mother and, while most will show only the premutation, others will express the full mutation. Sons who inherit an X chromosome from a mother carrying a premutation are the principal group affected by fragile X syndrome, since they are much more likely than females to have the gene amplified to the full mutation. The probability of the full mutation is dependent upon events in the mother; those at the lower end of the premutation range, about 56–70 repeats, are less likely to have a son with the full mutation than those with more than 100 repeats.

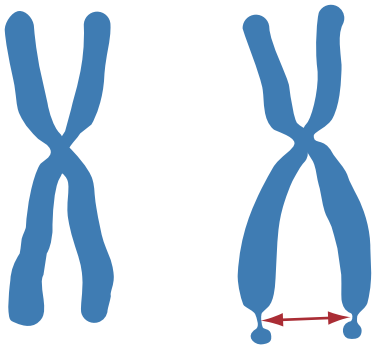


Figure 15.24 Schematic of a human X chromosome with a fragile site (arrow).

Margin Note 15.4
MicroRNA

MicroRNA (miRNA) is noncoding RNA with an average length of only 22 nucleotides. MicroRNA was first identified and shown to have a regulatory role in the nematode *Caenorhabditis elegans*. Since their discovery, it is estimated over 2300 human miRNAs occur. Most miRNAs are transcribed from DNA sequences, giving primary miRNAs, which are then processed to form mature miRNAs. In most cases, miRNAs interact with the 3′ untranslated regions (UTRs) of target mRNAs and suppress their expression. The discovery of miRNA has revolutionised the understanding of the post-transcriptional regulation of gene expression.

BOX 15.2 DYNAMIC MUTATIONS AND DNA METHYLATION

In the early 1990s, types of mutation called dynamic or expansion mutations were identified that were associated with a number of genetic disorders that increased in severity or had an earlier onset over several generations. These disorders are referred to as trinucleotide repeat disorders. Repeating combinations of three nucleotides occur commonly in DNA molecules. All combinations of triplets, such as CGG, CAG, AGG and ACC, are found, but the first two are most common. Such sequences seem to be a normal part of DNA and are thought to have regulatory roles in gene expression. However, if the number of repeats becomes too large, an identifiable disease may be triggered. Fragile X syndrome is one such disease and is associated with an increase in the number of CGG repeats that are normally found in *FMR1*. One consequence of the expansion in the number of CGG repeats is that methylation of the regulatory region of the gene occurs and prevents the cell from expressing *FMR1* and synthesising fragile X mental retardation protein (FMRP), an RNA binding protein expressed in the brain. The lack of FMRP leads to fragile X syndrome.

Methylation of DNA is the addition of methyl ($-CH_3$) groups (*Chapter 10*) to some of its bases from the donor molecule *S* adenosylmethionine (SAM). Only a few per cent of A and C nucleotides are methylated and, in vertebrate eukaryotic cells, only the formation of 5-methylcytosine (Me^5C) occurs (*Figure 15.25*). Methylation at particular CpG sequences may inhibit transcription and have a role in gene regulation by switching off the expression of that gene, although the mechanism by which this occurs is unclear. However, the methyl group is known to project into the DNA molecule and interfere with the attachment of DNA binding proteins. The patterns of methylation are inherited, that is, they are repeated from generation to generation. In general, *FMR1* is copied by the DNA polymerase during DNA replication with high fidelity and this stable version is almost always inherited. However, in some circumstances amplification occurs that increases the number of CGG repeats in the daughter chromosome. This is possible because repeats containing G and C nucleotides can base pair with themselves to form hairpin structures (*Figure 15.26*), which increases the risk of slipped mispairing occurring during DNA replication leading to an increase in the number of triplet repeats.

The number of CGG repeats in the *FMR1* is the major factor that determines the presence or absence of fragile X syndrome. In fragile X syndrome, the mutant *FMR1* gene is nonmethylated in asymptomatic males, methylated in the inactive X chromosome of females and totally methylated in most fragile X males, which prevents its expression.

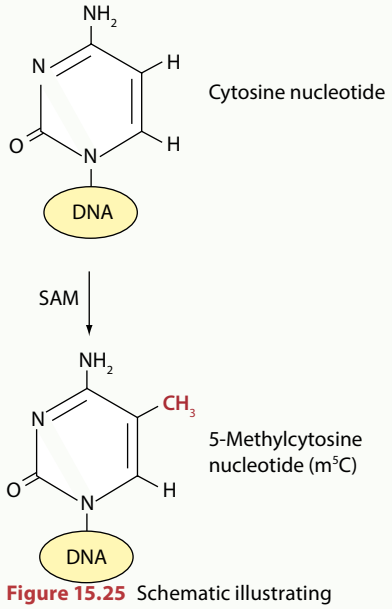


Figure 15.25 Schematic illustrating the methylation of a cytosine base in DNA forming 5-methylcytosine. SAM, *S* adenosylmethionine.

All males with the full mutation will experience significant symptoms. Some females with the full mutation will have symptoms of fragile X but, in general, the severity is less. Finally, there are individuals who cannot be assigned to these categories but have cells that vary regarding repeat size or the extent of methylation (*Box 15.2*). The severity of their symptoms depends on the proportion of cells affected and the tissues involved.

The phenomenon of trinucleotide repeats is seen in several other human disorders. For example, a fragile site on chromosome 3 containing the gene *FHIT* (fragile histidine triad) is often altered in cells from tumours of patients with lung cancer (*Chapter 17*). Huntington disease, myotonic dystrophy and spinobulbar muscular atrophy or Kennedy disease are also associated with trinucleotide amplifications, although they differ from fragile X syndrome in that the amplification can occur in both sexes at each generation

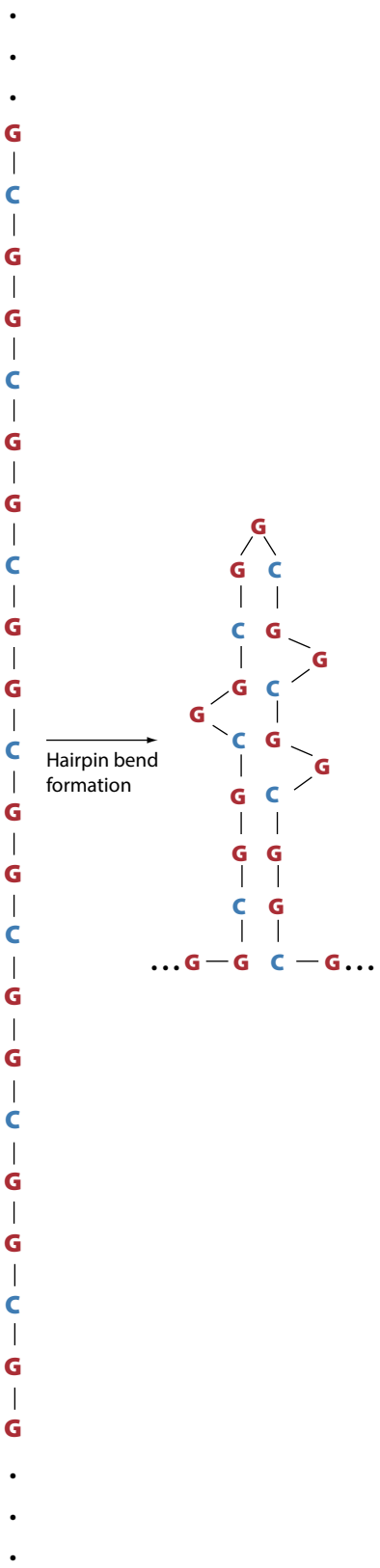


Figure 15.26 A highly schematic representation of a hairpin bend formed by a series of repeated CGG trinucleotides.

and is not associated with chromosome fragility. However, they are similar in that a threshold number of triplet repeats must be exceeded before symptoms of the disease appear.

15.8 CHROMOSOMAL MUTATIONS

Chromosomal mutations, sometimes called aberrations, include structural changes within a chromosome or changes in the numbers, including complete sets, of chromosomes.

Changes (mutations) in chromosome structure

Structural mutations occur when chromosomes break and, although in general repair mechanisms rejoin the two ends to rapidly restore the original structure, if more than one break occurs, the repair mechanisms are unable to distinguish between the broken ends, and portions of different chromosomes may be joined together. This can lead to one of four major types of chromosomal structural mutation:

1. deletions
2. duplications
3. inversions
4. translocations

In deletions, a section of the DNA is lost; in duplications, one or more extra copies of a segment of DNA occur in a chromosome. Inversions involve a reversal of direction of a portion of the DNA in the chromosome, while, in translocations, segments of the DNA are moved to a different chromosome. Deletions and duplications change the amounts of DNA in a chromosome. Inversions and translocations change the arrangement of bases in a length of DNA but do not change the amount of DNA present in the chromosome. If the exchange of chromosomal material does not involve breaks within a gene or alter the amount of DNA, then the individual will be clinically normal and is said to have a balanced translocation. However, if the structural alteration occurs in the gonads, even a balanced translocation has clinical significance for future generations, since it may lead to offspring who are chromosomally unbalanced, that is, who have lost DNA (*Figure 15.27*). Since most fetuses with unbalanced translocations tend to be spontaneously aborted,

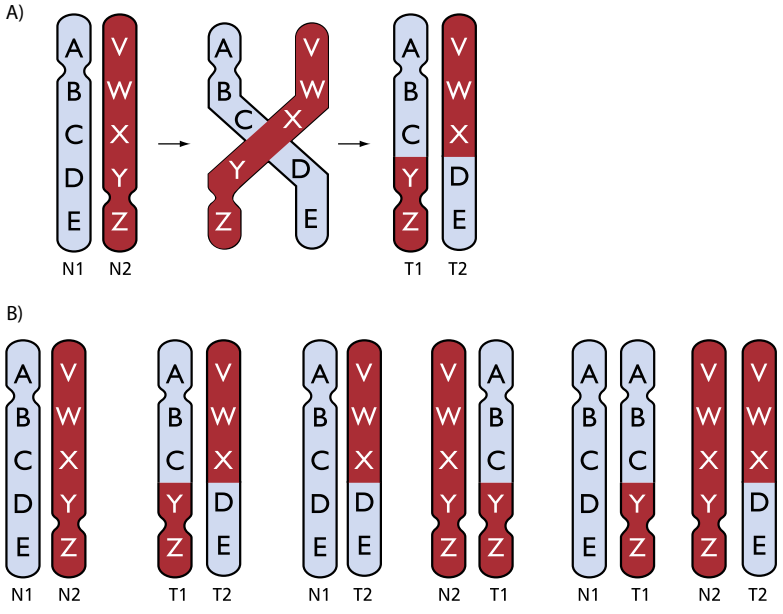


Figure 15.27 (A) Schematic showing a possible balanced translocation between two chromosomes. (B) Illustrates the types of gametes, both normal and genetically deficient, that could result from translocations shown in (A). For simplicity, it has been assumed that recombination did not occur during meiosis.

people with balanced translocations may only attract clinical attention if they and their partner are investigated because of a history of miscarriages. Infants with unbalanced translocations that do survive have reduced mental abilities and show multiple **dysmorphic** features, that is, alterations (abnormalities) to the accepted appearance. Specific disorders can also occur when discrete genes are damaged at the translocation fractures, the resulting disorder being dependent upon which gene(s) is(are) damaged.

Deletions

Deletion of part of a chromosome can arise between two breakpoints because of unequal crossing over during meiosis (*Section 15.3*) or from a parental translocation.

Several clinical disorders are caused by deletions but, in many cases, abnormalities only occur in heterozygous individuals, since the homozygous condition is lethal, especially in cases involving large chromosomes. Many chromosomal deletions that can be observed microscopically almost invariably produce a phenotype with multiple abnormal features and reduced mental facilities because of the absence of expression of the deleted genes. For example, cri-du-chat syndrome is a heterozygous condition that occurs in about 1 in 50 000 births. Infants with this syndrome show anatomical malformations, including gastrointestinal and cardiac complications, and are often mentally affected. The glottis and larynx also develop abnormally, giving the characteristic cry, similar to the meowing of a cat, which names the syndrome. Cri-du-chat syndrome is caused by a deletion of part of the short arm of chromosome 5 (46,5p-; *Figure 15.28*). The length of the deleted portion varies: the longer the deletion, the more defective the syndrome in surviving children. The effects of the syndrome are severe, although individuals who receive good home care and early special schooling can walk and communicate verbally and develop self-care skills.

Prader–Willi syndrome (frequency: 1 in 10 000–30 000 births) and Angelman syndrome (frequency: 1 on 15 000) both result from the 'microdeletion' of the band 15q11-13, although they have different phenotypes depending upon the parental origin of the deletion. In Prader–Willi syndrome, the deletion occurs invariably on the chromosome 15 inherited from the father, whereas in Angelman syndrome the deletion occurs almost exclusively on the chromosome 15 from the mother. Hence, inheritance of paternal and maternal copies of this region of chromosome 15 is important for normal development, a phenomenon known as **genetic imprinting**. People with Prader–Willi syndrome have learning disabilities, have small external genitalia and characteristic facial features. Babies with Prader–Willi syndrome have a poor sucking reflex; hence, feeding is difficult and results in weakness and stunted growth. Strangely, children with Prader–Willi syndrome become compulsive eaters at five to six years of age and become obese (*Chapter 10*) and suffer the related health problems, such as type 2 diabetes (*Chapter 7*) and heart failure (*Chapter 14*).

Angelman syndrome is characterised by developmental delay, absence of speech, jerky movements, paroxysms of inappropriate laughter and characteristic facial features, which differ from those of Prader–Willi syndrome.

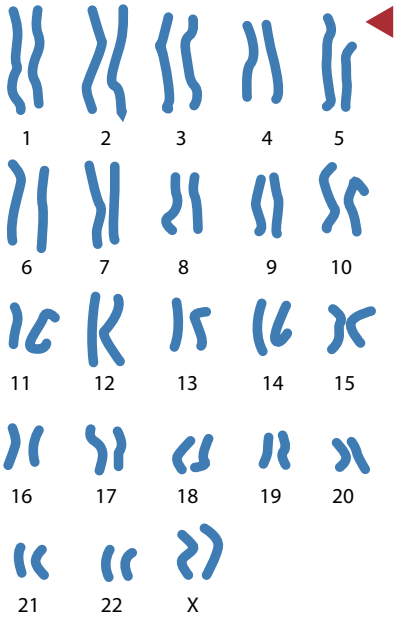


Figure 15.28 Deletion of part of a chromosome 5 (arrow) in the karyogram of a child showing cri-du-chat syndrome. ⏎

Duplications

Duplications are chromosomal mutations that result in the doubling of part of a chromosome. The size of the duplicated segment varies enormously, and duplicated segments may occur in a tandem configuration, that is, adjacent to each other, or in different locations in the genome (*Figure 15.29*).

Duplications can result in gene redundancy and may produce phenotypic variation, since there are now two copies of the gene, and one copy may mutate independently of the other. This is a significant source of potential genetic variability during evolution. For example, gene duplications have been essential in the evolution of multigene families. The genes for globins are a particularly well-studied multigene family. Haemoglobins are tetrameric proteins consisting of two pairs of differing polypeptides each with an attached haem group able to bind and release a dioxygen molecule. Human individuals produce different haemoglobins at different stages during their lives as described in *Chapter 13*. Each of the globin polypeptides has slightly different primary structures, but consist of the α types that are found as a cluster on chromosome 16 and the β types clustered on chromosome 11. It is thought that each group of genes evolved from one original ancestral gene that underwent duplications and subsequent sequence divergence. Serine proteases (*Chapter 11*), such as the digestive enzyme trypsin, are another well-documented family of proteins.

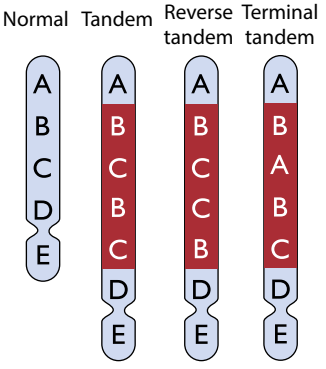


Figure 15.29 Schematic to illustrate some possible types of chromosomal duplications. ⏎

Inversions

Figure 15.30 illustrates how inversions might arise. A chromosomal loop forms before fractures occur at two places on the chromosome. The insertion of the inverted segment

at the newly created sticky ends and their subsequent joining within the chromosome completes the inversion. There are two types of inversions: paracentric inversions do not include the centromeres, whereas pericentric inversions do. Genetic material is not lost during inversions, although there can be clinical problems when fractures occur within genes or within regions that control gene expression. The meiotic consequences of a chromosomal inversion depend on the type of inversion encountered and the resulting gametes may be nonviable, leading to reduced fertility.

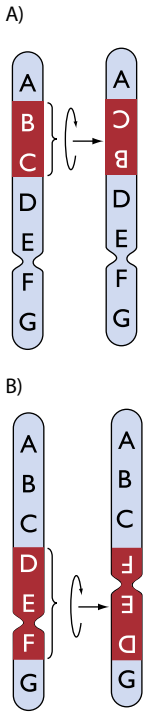


Figure 15.30 Schematics to show (A) a paracentric and (B) a pericentric inversion.

Translocations

Numerous translocations occur in the human population. The simplest kinds are intrachromosomal translocations that move part of a chromosome to a different position within the same chromosome. Interchromosomal translocations transfer part of a chromosome to a nonhomologous chromosome (*Figure 15.31* and *Box 15.3*). Reciprocal translocation involves an interchromosomal translocation between two nonhomologous chromosomes. The least complex way this event can occur is for two nonhomologous chromosome arms to come close to each other so that an exchange is facilitated.

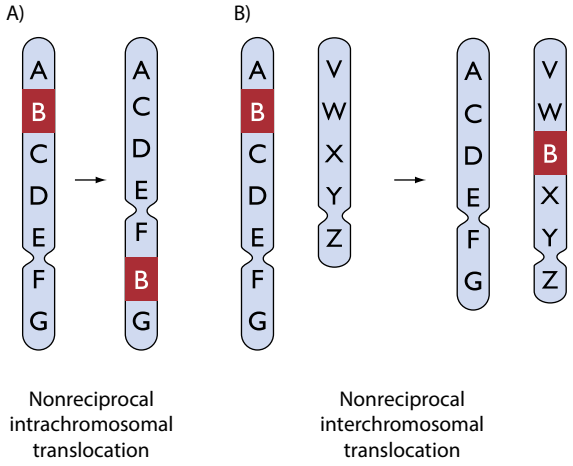


Figure 15.31 (A) Intrachromosomal and (B) interchromosomal translocations.

Homologues that are heterozygous for a reciprocal translocation undergo unorthodox synapsis during meiosis, and pairing results in mitotic figures with a cross-like configuration. These chromosomes produce genetically unbalanced gametes and often result in reduced fertility. As few as 50% of the progeny of parents that are heterozygous for a reciprocal translocation survive, a condition known as semisterility. In humans, translocation can result in variations from the normal diploid number of chromosomes leading to a variety of birth defects. Translocations may transfer a gene to a region of a chromosome that is more transcriptionally active. This can lead to the development of some forms of cancer, for example alveolar rhabdomyosarcoma, a muscle tissue tumour of children, which is linked to the translocation t(2,13) (q35,q14) that produces the hybrid protein PAX3–FOXO1 (*Chapter 17*).

BOX 15.3 ACUTE PROMYELOCYTIC LEUKAEMIA

Leukaemias are tumours that result in the overproduction of immature leukocytes in the bone marrow and their release into the circulation (*Chapter 17*). Several types of leukaemia have been linked with chromosomal disorders, including acute promyelocytic leukaemia (APL). The condition constitutes 80% of adult and 20% of childhood leukaemias.

Promyelocytic leukaemia protein (PML) is a **tumour suppressor protein** (*Chapter 17*) required for the assembly of structures in the nucleus called PML-nuclear bodies. These bodies have regulatory cellular functions, including involvement in **apoptosis** (*Chapter 17*) and controlling cell division. Mutations involving PML are implicated in several types of cancers. Some forms of vitamin A (*Chapter 10*), collectively called retinoids, are also involved in regulating cell proliferation and differentiation and therefore help control embryonic development and the differentiation of skin and other organs. All-*trans*-retinoic acid (ATRA) and its isomer 9-*cis*-retinoic acid act biologically in similar ways to steroid hormones (*Chapter 7*). They penetrate the plasma membrane of target cells and bind to intracellular receptor proteins. Two groups of receptors, termed retinoic acid receptors (RAR) and retinoid X receptors (RXR), have been identified, each of which is a subgroup of the nuclear receptor (NR) superfamily of transcription factors. The binding of retinoic acid leads to the formation of a RAR–RXR heterodimer (*Figure 15.32*) that is translocated to the nucleus where it interacts with specific DNA sequences, called retinoic acid response elements (RARE), located in the promoter regions of retinoid-regulated genes, which leads to alterations in the gene expression of the cell.

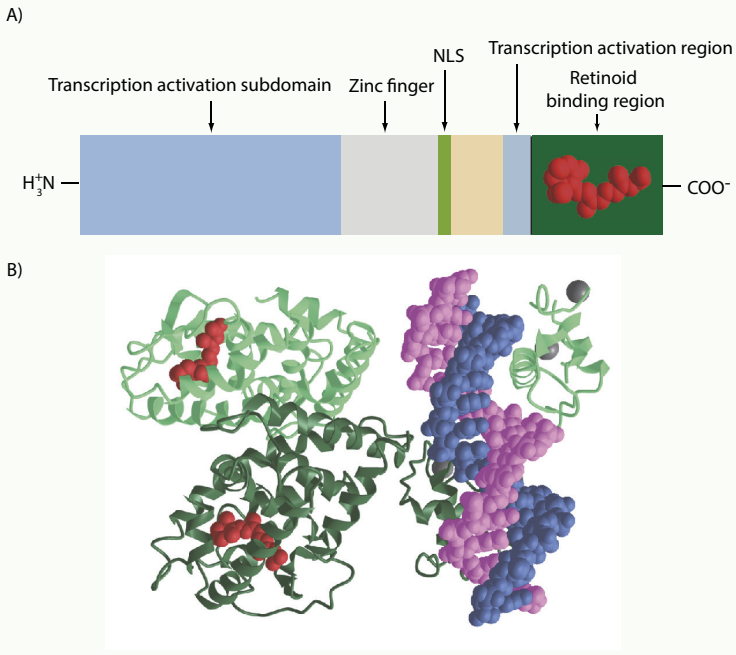


Figure 15.32 (A) Cartoon showing the generalised structure of a retinoid receptor. The nuclear locating signal (NLS) is a sequence of basic amino acid residues that ensures that the complex enters the nucleus. The zinc finger motifs are responsible for binding the complex to the DNA at retinoic acid response elements (RARE), located in retinoid-regulated genes. The amino and carboxy termini contribute to activating transcription and binding the retinoid, respectively. (B) Molecular model of a retinoid–receptor complex bound to DNA. The retinoic acid (RAR) and retinoid X (RXR) receptors are shown in darker and lighter greens, respectively, while the bound retinoic acids are highlighted in red. The DNA is shown in shades of magenta. The Zn^{2+} that stabilise the zinc finger binding motifs are indicated as grey spheres. PDB file 5UAN.

In APL, there are abnormal hypergranular promyelocytes or immature granulocytes (*Chapter 13*) and the bleeding disorder called disseminated intravascular coagulation (DIC) occurs. Patients may present with bruising and severe bleeding, which tends to worsen when treatment is started, as the leukaemic cells break down and consume large amounts of clotting factors and platelets, and there is also an increased risk of infections. The chromosomal defect that causes APL is a translocation between parts of the long arms of chromosomes 15 and 17, which is balanced and reciprocal, and results in one abnormally long chromosome 15 (15q+) and a shortened chromosome 17 (17q–). Clinically, this is significant because PML is located at 15q24.1 and encodes the tumour suppressor PML (*Chapter 17*), and RAR, which encodes a retinoid receptor, is found at 17q21.2. The protein product of the fused oncogene, *PML-RAR*, interferes with the normal functioning of PML and with the roles of RAR in myeloid differentiation to produce different types of blood cells from progenitor cells in the bone marrow and results in APL A diagnosis of APL is suggested by the general clinical picture, examining blood smears and bone marrow aspirates, and cytological examinations. Treatment with ATRA is beneficial, especially when combined with conventional chemotherapy. Patients must be closely monitored for DIC and clinical signs, such as fever, oedema in the limbs, and kidney, liver and respiratory dysfunctions.

A common type of translocation involves breaks at the extreme ends of the short arms of two nonhomologous acrocentric chromosomes (*Figures 15.33* and *15.36*). The smaller fragments are lost, but the larger ones fuse together at their centromeric regions. This type of translocation produces a new, large submetacentric or metacentric chromosome and is often called a Robertsonian translocation.

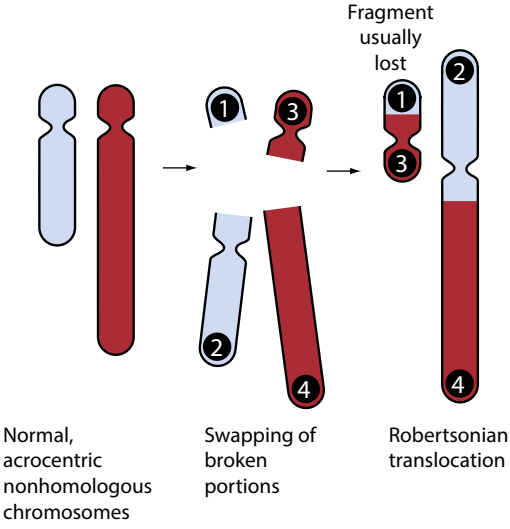


Figure 15.33 The formation of a Robertsonian translocation following breaks in two acrocentric chromosomes.

Variations in chromosome numbers

Eukaryotic organisms are normally diploid and produce haploid gametes (*Section 15.3*). However, mutations with numerical aberrations of the chromosome complement, including whole sets, do occur. *Table 15.7* lists several human chromosomal abnormalities. These types of mutations are divided into two major types:

1. aneuploidy
2. euploidy

Chromosomes	Syndrome
Autosomes	
Trisomy 13	Patau syndrome
Trisomy 18	Edwards syndrome
Trisomy 21	Down syndrome
Sex chromosomes, female	
X0	Turner syndrome
XXX	Triple X syndrome (trisomy X)
XXXX	Tetrasomy
Sex chromosomes, male	
Y0	Nonviable
XYY	XYY syndrome
XXY, XXXY, XXYY	Klinefelter syndrome

Table 15.7 Aneuploid abnormalities of human chromosomes. ⏎

Aneuploidy

Aneuploidy is usually caused by the nondisjunction of paired chromosomes at meiosis I or of sister chromatids at meiosis II or by delayed movement of a chromosome at anaphase. Nondisjunction results from the failure of pairs of homologues to separate or disjoin during segregation. *Figure 15.34* illustrates the consequences of nondisjunction during first meiosis and second meiosis for a single chromosome. Thus, gametes are formed that either lack the chromosome or contain two copies of it. If these are fertilised by a normal haploid gamete, then zygotes are produced containing one or three of that chromosome. Aneuploidy has serious, often lethal, consequences in humans. Indeed, approximately 90% of all chromosomal aberrations induce a spontaneous abortion.

Monosomy is the loss of a single chromosome from an otherwise diploid genome, that is, 2N–1. Nullisomy results from the loss of one pair of homologous chromosomes (2N–2). The gain of one chromosome results in trisomy (2N+1). Tetrasomy describes the presence of four copies of a specific chromosome rather than the normal two (2N+2). Aneuploidy can also involve the loss or the addition of more than one particular chromosome or pair of chromosomes. Thus, a double monosomy involves the loss of two separate nonhomologous chromosomes (2N–1–1), while a double tetrasomy would describe the presence of two copies of two different chromosomes (2N+2+2). Both of these cases involve meiotic nondisjunction in two different chromosomes prior to gamete formation. Monosomy results in two types of haploid gametes: N and N–1. Also, the single unpaired chromosome in the 2N–1 cell is easily lost during meiosis, resulting in the formation of two gametes with N–1 chromosomes.

Autosomal monosomy is rare, since monosomic embryos do not develop significantly and are usually lost early in pregnancy, although the abnormalities can be detected in aborted fetuses. In trisomy, the individual concerned has more chance of surviving if the extra chromosome involved is relatively small. The addition of a large autosomal chromosome has severe effects and is usually lethal during development. Autosomal trisomies are found in about half of the chromosomal abnormalities that lead to fetal death.

Trisomy 13 (47,13+) produces Patau syndrome (*Figure 15.35*) and affects about 1 in 5000 live births. The syndrome is characterised by many abnormalities, including poor

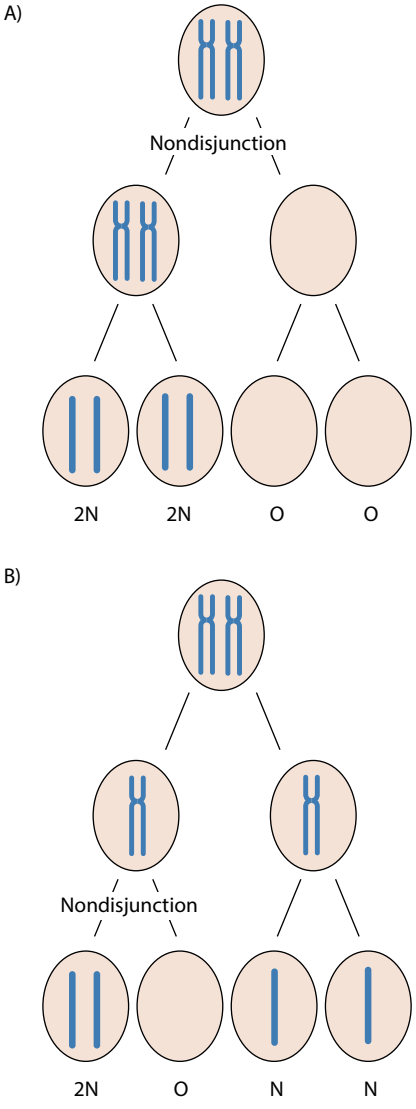


Figure 15.34 The consequences of nondisjunction of a single chromosome at (A) meiosis I and (B) meiosis II. ⏎

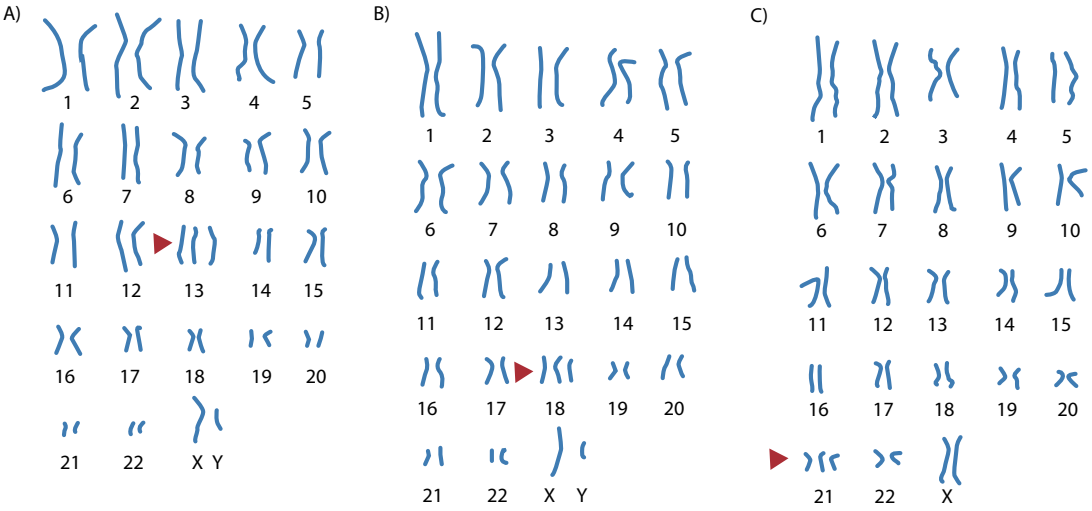


Figure 15.35 Karyotypes showing (A) trisomy 13, Patau syndrome, (B) trisomy 18, Edwards syndrome, and (C) trisomy 21, Down syndrome.

mental and physical development, cardiac anomalies, polydactyly (that is, extra fingers or toes), cleft lip and palate, and small eyes. Approximately 90% of affected babies die before they are one year old.

Trisomy 18 (47,18+) is associated with Edwards syndrome (*Figure 15.35*) and occurs in about 1 in 4000 live births. About 80% of cases of are female. Sufferers are small at birth and have multiple congenital malformations that affect almost all body systems. Among the many abnormalities associated with the syndrome are restricted mental and body development, elongated skull, low-set malformed ears and clenched fists. Ninety per cent of infants with trisomy 18 die within six months, often from cardiac problems.

Trisomy 21 (*Figure 15.35*) leads to Down syndrome, the only human autosomal trisomy in which significant numbers of individuals survive more than a year following birth. It was named after its 'discoverer', a doctor called Langdon Down (1828–1896), in 1866. Down syndrome affects, on average, 1 in every 700 live births, with the incidence increasing with the age of the mother but not that of the father (*Figure 15.36*). However, in terms of gross numbers, most affected children are born to women under 35 years old, since most pregnancies occur below this age. The relationship between the age of the mother and the incidence of Down syndrome is explainable in terms of the production of oocytes (*Chapter 7*). Females have a full complement of primary oocytes that developed in the ovary of the developing female fetus. These oocytes have commenced meiosis but are arrested at prophase I before birth. In a fertile adult female, the nucleus of a secondary oocyte begins the second meiotic division at each monthly ovulation but progresses only to metaphase II, when division again stops. The second meiotic division is not completed unless a sperm penetrates the secondary oocyte; hence, the succeeding ovum has been arrested in meiosis for about a month longer than the preceding one. Older ovulating women produce ova that are significantly older and have been arrested in meiosis longer than those of younger women. It is possible that the probability of nondisjunction increases with the length of time the primary oocyte spends in the ovary (*Chapter 7*), but Down syndrome can originate from either parent; approximately 5% of cases have been traced to the father.

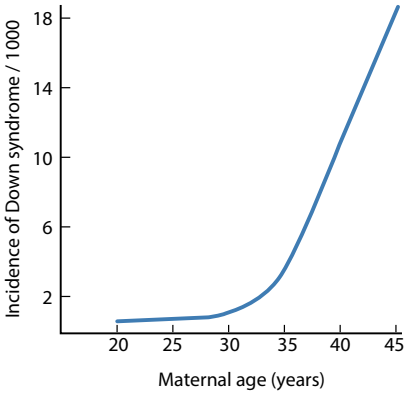


Figure 15.36 The effect of increasing maternal age on the incidence of Down syndrome babies.

Over 95% of all cases of Down syndrome are caused by trisomy 21 (*Figure 15.35*) following a nondisjunction event during meiosis. Most people with Down syndrome have 47 chromosomes in all their cells. Approximately 3–4% have the normal number (46) but have a type of Down syndrome that runs in families called familial Down syndrome. This results from a Robertsonian translocation where the long arm of chromosome 21 joins the long arm of chromosome 14 or sometimes 15, producing three copies of the long arm of chromosome 21. The heterozygous carrier is normal because there are two copies of all major chromosome arms and hence two copies of all essential genes (*Figure 15.37*). However, meiosis will result in a 25% of the gametes formed having two copies of chromosome 21: one normal chromosome 21 and a copy attached to chromosome 14. When this gamete is fertilised by a normal haploid gamete, it leads to Down syndrome because, although the zygote has the normal 46 chromosomes, it has three copies of chromosome 21.

Affected individuals have common physical features and are often described as having affectionate, loving natures. They generally have flat faces with epicanthic folds over

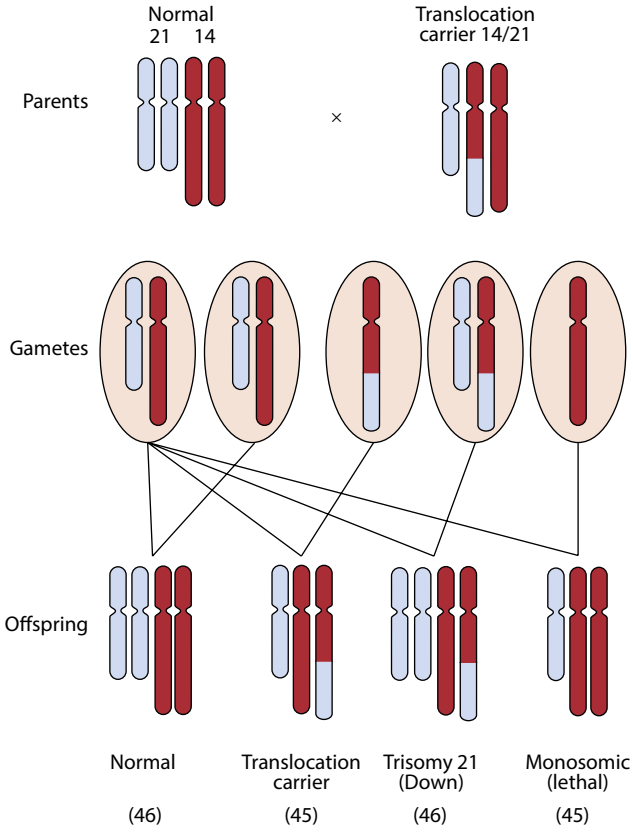


Figure 15.37 A Robertsonian translocation involving chromosomes 14 and 21 resulting in trisomy 21, Down syndrome.

the eyes and round heads with a protruding furrowed tongue that causes the mouth to remain partially open. They are below average height and have short, broad hands. Their physical and mental development is restricted and muscle tone and motor skills are poor. Down syndrome patients are prone to respiratory disease, 50% of them have heart problems and their incidence of leukaemia is approximately 15 times higher than that of the normal population. Chromosome 21 is known to carry *APP*, a gene that encodes the amyloid β precursor protein associated with Alzheimer disease (*Chapter 18*). About 30% of adults with Down syndrome develop Alzheimer disease at 50–60 years of age, and this increases to over 50% for those in their 60s.

Some individuals with Down syndrome must be institutionalised, but most can be cared for at home and benefit greatly from special education programmes. Advances in several areas of clinical management have, fortunately, resulted in people with Down syndrome living increasingly longer lives with an improved wellbeing.

Aneuploidy involving sex chromosomes

Aneuploidy involving sex chromosomes gives rise to several well-defined syndromes, for example:

- Turner
- Klinefelter
- Jacob or XYY

A female born with a single X chromosome (45,X0) shows Turner syndrome (*Table 15.7*). The incidence is 1 per 5000 female births, although most 45,X0 aneuploids are spontaneously aborted. The condition may arise from nondisjunction in either parent but, in 75% of cases of Turner syndrome, only the maternal X chromosome is present, implying that the problem originates in spermatogenesis in the father. Females with Turner syndrome are short in stature and rarely undergo secondary sexual development and so are mostly infertile, although their intelligence and lifespan are generally normal. Symptoms may be reduced with oestrogen and growth hormone treatment.

A male born with one or more extra X chromosomes (*Table 15.7*) exhibits Klinefelter syndrome. The incidence of 47 XXY is 1 in every 1000 boys, although the risk increases with an increase in the age of the mother. The extra chromosome is donated from the

mother in 60% and from the father in 40% of cases and arises by nondisjunction during both maternal and paternal meiotic divisions. Variants include XXYY, XXXY and even XXXXY. Affected males have underdeveloped testes, are infertile, are often above average height and have mild mental disabilities. Long-term treatment with testosterone can improve bone and muscle mass and, while sustaining sexual function, does not restore fertility.

Approximately 1 in 1000 male children exhibit XYY or Jacob syndrome. It arises from a nondisjunction of the Y chromosome. Males with XYY syndrome are above average in height and may be less fertile, but often only come to medical attention because of problems with fine motor coordination with low muscle tone (hypotonia), speech disorders and learning disabilities. However, in others with the syndrome, the signs and symptoms may be barely noticeable and they do not exhibit significant learning difficulties. Management is to treat symptoms and offer support, which may include the need for special education. Occupational therapy may be recommended for infants and young boys who have hypotonia, and speech therapy may be necessary when speech is delayed. However, the long-term prognosis is generally favourable for a productive civil life.

Euploidy

Monoploidy is the presence of only a single set (that is, 23) of chromosomes rather than the normal diploid 46. Monoploid fetuses do not reach full term, presumably because the recessive lethal mutations, which are usually counteracted by dominant alleles in heterozygous individuals, are expressed. In **polyploidy**, the chromosome number is an exact multiple of the haploid number but exceeds the diploid number. Organisms with three sets of chromosomes are triploid (*Figure 15.38*), those with four sets are tetraploid and so on.

Polyploidy usually arises from fertilisation of the egg by two spermatozoa, which increases the total number of chromosomes to 69, or from a failure at one of the maturation divisions of the egg or spermatozoon so that a diploid gamete is produced. If nondisjunction occurs at meiosis I, 50% of the gametes lack chromosomes and the other 50% have two sets of chromosomes. If nondisjunction occurs at meiosis II, 50% of the gametes possess the normal single set of chromosomes, 25% have two sets of chromosomes and 25% lack chromosomes (*Figure 15.39*).

Fusion of a gamete with two chromosome sets with a normal gamete produces a zygote with a triploid set of chromosomes (3N). Similarly, fusion of two gametes, each with two chromosome sets, produces a tetraploid (4N) zygote. Polyploidy of somatic cells can also occur following the mitotic nondisjunction of complete chromosome sets.

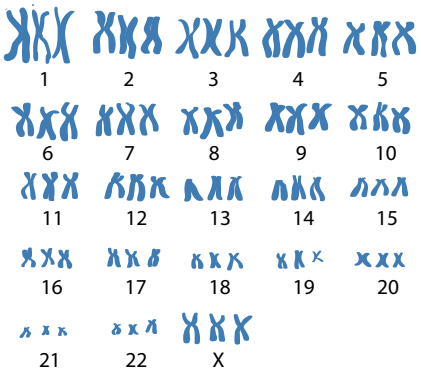


Figure 15.38 A triploid karyotype. Courtesy of J.S. Haslam and K.R. O'Craft, Tameside General Hospital, Ashton under Lyne, UK.

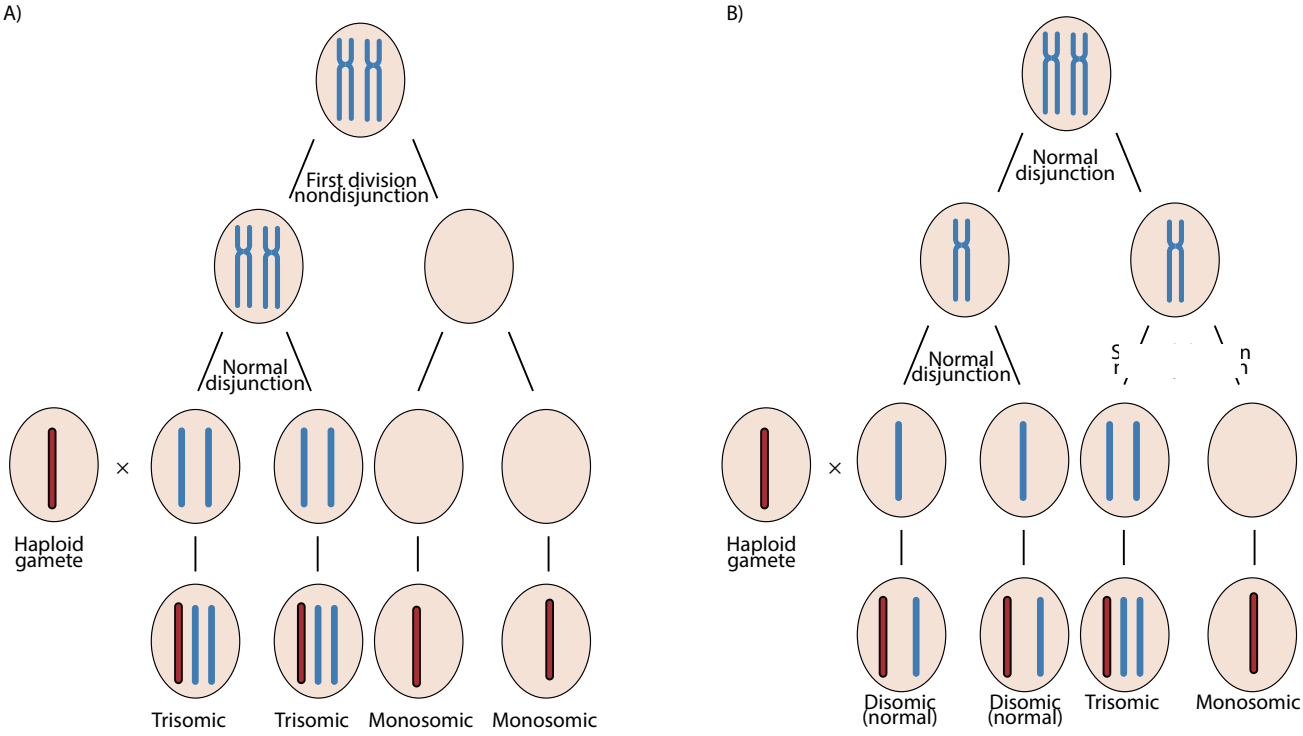


Figure 15.39 Schematic showing the production of abnormal gametes following nondisjunction at (A) meiosis I and (B) meiosis II that result in trisomic zygotes following fertilisation with normal haploid gametes.

Polyploidy normally causes an early spontaneous abortion and survival of the fetus to full term is rare. The most common type of polyploidy in humans is triploidy. However, polyploids with an odd number of chromosome sets always possess an unpaired chromosome for each type; hence, the probability of producing a balanced gamete is low. Indeed, triploidy in humans is always lethal and is seen in 15–20% of spontaneous abortions and in only about 1 in 10 000 live births, where death almost always occurs within a month. Triploid babies show many abnormalities, including a characteristically enlarged head. Polyploids with even numbers of sets of chromosomes generally have a better chance of being at least partially fertile because there is the potential for homologous chromosomes to be segregated equally during meiosis. However, tetraploidy is also always lethal in humans and is seen in approximately 5% of spontaneous abortions. Very rarely, tetraploid humans are born, although they only survive for a short time.

15.9 DETECTING, DIAGNOSING AND SCREENING HUMAN GENETIC DISEASES

Identifying the genetic basis for a human disorder usually requires an examination of the family history as far back as possible. If this indicates that the trait is inherited, it is possible to predict whether the mutant allele is dominant or recessive and whether it is X- or autosomal-linked. Obviously, dominant mutations are the simplest to detect. If these are X-linked, then affected fathers will pass the trait to all their daughters. However, if autosomal, then dominant mutations would be expected to occur in approximately 50% of the children of an affected heterozygous individual.

In many cases, the diagnosis of a genetic disease can be made prenatally. The most widely used methods for obtaining sample material are chorionic villus sampling (CVS) and amniocentesis. The former is normally performed between 11 and 14 weeks of pregnancy, while amniocentesis is usually offered from 15 weeks. In CVS, a catheter is inserted into the uterus and a small tissue sample of the fetal chorion removed. Amniocentesis involves using a syringe to withdraw amniotic fluid, which contains cells genetically identical to those of the fetus (*Figure 15.40*). In both cases, a variety of cytogenetic and biochemical tests can then be performed on the cells and tissue to identify any single gene disorders or chromosomal abnormalities.

Margin Note 15.5 Noninvasive prenatal testing

Noninvasive prenatal testing (NIPT) is based on the detection and characterisation of circulating cell-free fetal DNA in maternal plasma with the aim of identifying genetic diseases. In 1997, the analysis of cell-free DNA (cfDNA) was first used to identify the presence of Y chromosomal DNA circulating in the blood of pregnant women carrying male fetuses. Cell-free fetal DNA fragments originate from placental trophoblastic cells released into the maternal circulation during pregnancy and are representative of the entire fetal genome. It is detectable from as early as four weeks of gestation. Noninvasive prenatal testing is now offered within the national screening pathway for Down, Edwards and Patau syndromes.

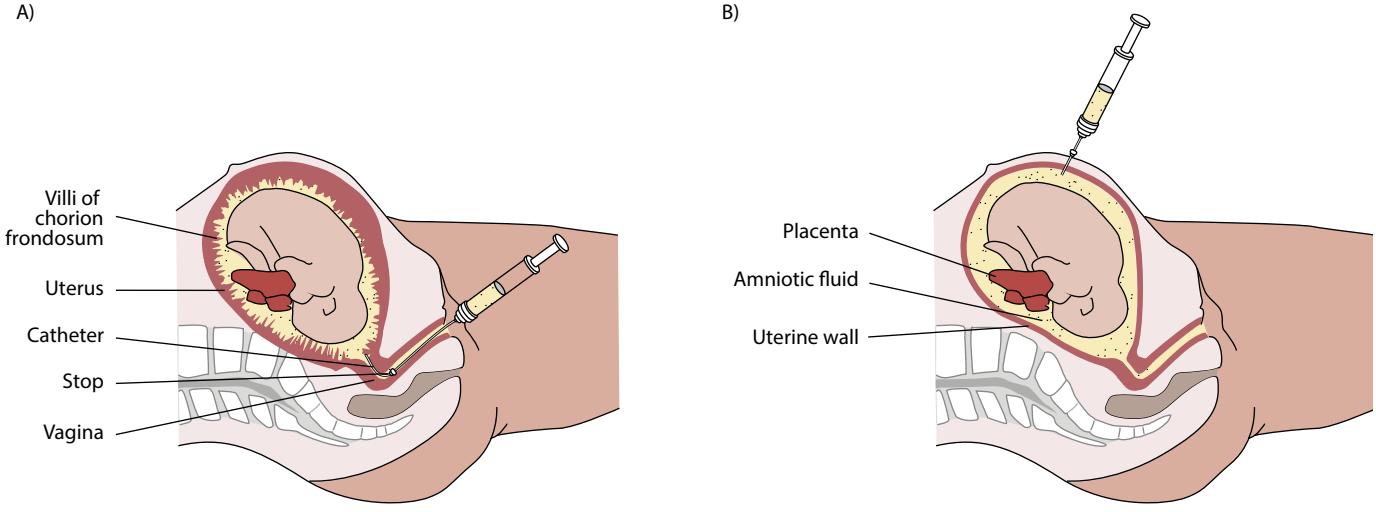


Figure 15.40 Outlines of (A) chorionic villus sampling and (B) amniocentesis to obtain specimens for genetic analysis of the fetus.

People in whom a genetic defect has been detected will be advised to seek genetic counselling, especially if they wish to have children. The advice they are given is based on analyses of the risks that they may produce a child with a genetic abnormality or that they themselves may develop a late-onset genetic disease. A broad range of information is required, and genetic counsellors require training to provide this information with appropriate consideration.

Detection of hereditary diseases

Karyotyping is a major method of identifying a hereditary disorder, such as Down syndrome (*Box 15.4*). Tests based on recombinant DNA technology have increased in recent years because of their high sensitivity and accuracy in detecting genetic disorders at the

BOX 15.4 PRENATAL DIAGNOSIS OF DOWN SYNDROME ⏎

The quadruple test for Down syndrome calculates the chance of having a fetus with Down syndrome, taking into account the maternal age at term and four markers (α fetoprotein [AFP], unconjugated oestriol, total human chorionic gonadotropin [hCG; or more usually its free β subunit] and inhibin-A) at 14–22 weeks, in the maternal serum. There have been several variations on this test. In Down syndrome, the AFP tends to be lower than normal concentrations (while, in neural tube defects, for example spina bifida, it is higher). However, concentrations of hCG are increased. A screening protocol was proposed in 1998 that combined maternal serum AFP, the urinary β-core fragment of hCG (a breakdown product of hCG) and total urine oestradiol. This test is said to be superior to the 75% sensitivity at the 5% false-positive (*Chapter 1*) rate compared with when β-core/oestradiol and maternal age alone are used.

The definitive test for Down syndrome is prenatal cytogenetic screening, that is, by karyotyping (*Section 15.7*). Chorionic villus sampling (CVS) or amniocentesis and the culture of fetal cells will allow its karyotype to be determined. In the developed countries, this is routinely offered to pregnant women. Karyotyping detects a range of numerical and structural abnormalities in the chromosomes in addition to Down syndrome and other common autosomal trisomies. However, the cells obtained from amniocentesis or CVS must be cultured to provide sufficient biological material, and so full cytogenetic analysis involves a delay of 14 days or longer before a result can be given.

The statistics relating the risk of bearing a Down syndrome child to the age of the mother (*Figure 15.36*) raise serious issues for women who become pregnant later in life. Genetic counselling may also be useful in informing the parents about the probability that their child will be affected and in educating them about Down syndrome. Older pregnant women should consider tests to determine whether the fetus has a normal complement of chromosomes. Obviously the earlier a definitive positive result is obtained, the sooner the prospective parents can make clinically relevant decisions. Therefore, 'molecular methods' have been investigated, since these have the potential to give a result much more rapidly.

Fluorescence *in situ* hybridisation (FISH) is a technique in which fluorescently labelled DNA probes that will hybridise to relevant regions of the chromosomes are applied to cell preparations (*Chapter 5*). When the preparations are examined with a fluorescence microscope, the presence or absence of the target regions are revealed. Alternatively, the polymerase chain reaction (PCR; *Chapter 3*) can be used to amplify these regions and the results can be seen on electrophoresis gels. Both methods can potentially give a result in 24–48 hours. These molecular methods have been shown to be accurate, since they would fail to detect an abnormal karyotype in only about 1 in 100 amniocentesis samples. Obviously, such tests will only give information about regions of the chromosomes for which DNA probes are applied. Many clinical laboratories in the UK now offer FISH or the PCR along with karyotyping, mostly to target pregnancies that are considered to be at high risk after screening by the quadruple test. If the fetus is diagnosed as having Down syndrome, a therapeutic abortion is an option the parents may consider.

prenatal stage. The detection of hereditary diseases in a given genome was traditionally a long and laborious task. However, modern techniques of chromosome and DNA analyses have resulted in a dramatic increase in the number of tests for genetic diseases, with a considerable improvement in the time required for analysis. For example, using the polymerase chain reaction (PCR; *Chapter 3*), virtually any portion of a gene or even whole genes can be amplified for analysis by electrophoretic techniques or sequenced to detect mutations. The use of next-generation sequencing methods (*Chapter 17*), which are rapid and inexpensive, has greatly facilitated diagnosis.

These molecular techniques allow not only diagnoses in patients with inherited disorders, but also the detection of mutations in carriers, even though they do not express any symptom of the disorder. They are also direct ways of distinguishing between different mutations within a single gene, where each mutation could lead to a different disorder, as seen, for example, in muscular dystrophies (*Chapter 16*).

Predictive tests for disorders that have only a partial genetic component, such as heart disease and cancers, have been developed. Indeed, the PCR analysis of cells lost in faeces has demonstrated premalignant changes in the gastrointestinal tract and allowed patients at risk of developing colon cancer to be identified (*Chapter 17*). This is of clinical importance, since the earlier that treatments of malignant conditions are started, the more favourable the prognosis.

Cloned DNA sequences have expanded the range of prenatal testing because they allow the fetal genotype to be examined directly, rather than relying on secondary tests for the products of the normal or mutant genes. Thus, mutations in DNA can be detected in those cases where an aberrant product cannot be detected prior to birth, even though a test is available. A test for sickle cell anaemia using a restriction enzyme is described in *Chapter 13*. However, restriction fragment length polymorphism (RFLP) analysis (see, for example *Chapter 13*) requires that the mutation in the gene alters the restriction sites recognised by restriction endonucleases and this is not always the case. In contrast, allele-specific oligonucleotides (ASOs) are synthetic nucleotide probes, which bind only to their complementary DNA. They will not hybridise to other sequences and, in appropriate highly stringent conditions, they differentially bind to and distinguish between alleles that differ by as little as one nucleotide. They can thus distinguish between the native and mutant forms of a gene with excellent resolution and potential versatility.

Tests based on ASOs are now available to screen for mutations associated with cystic fibrosis (*Chapter 16*) and glaucoma.

Potential parents in whom a genetic defect is detected must be advised to have genetic counselling. The advice they are given is based on analyses of the risks of them having a child with a genetic abnormality or that they may develop a late-onset genetic disease. Information drawn from a wide range of sources is required in such circumstances, and genetic counsellors must provide such information with tact and sensitivity.

Newborn screening tests

In the UK, the parents of all newborn babies are offered screening tests to examine their baby(ies) for any potential health problem. The tests consist of a physical examination three days after birth of the eyes, heart, hips and, in the case of boys, testicles. A follow-up examination is performed at six to eight weeks of age, given that some clinical conditions may develop after three days. For hospital births, a hearing test may be done before the mother and baby are discharged, otherwise it will be performed by a health visitor or other healthcare professional within the four or five weeks following birth, but it can be done up to three months of age. Genetic/inherited clinical conditions are tested for with the newborn blood spot or heel prick test. This involves taking a small sample of blood, usually at about five days or, less often, up to eight days of age. It involves pricking the skin on the heel and collecting drops of blood onto a blood spot card. The card is then analysed to detect if the mutant genes responsible for 9 or 10 rare clinical conditions are present (*Table 15.8*). Many developed countries test babies for a greater number of genetic conditions than the UK. The heel prick test is invasive and subjects the baby to some discomfort, but the obvious benefits of screening are early treatment to affected babies, which can improve health, prevent severe disabilities in some cases or even death, and allow support and appropriate advice to be given to the parents.

Genetic/inherited condition
Cystic fibrosis (*Chapter 16*)
Congenital hypothyroidism (*Chapter 7*)
Glutaric aciduria type 1 (GA1; *Box 9.2*)
Isovaleric acidaemia (IVA; *Box 9.2*)
Homocystinuria (pyridoxine unresponsive, HCU)
Maple syrup urine disease (MSUD)
Medium length fatty acyl-CoA dehydrogenase deficiency (MCADD)
Severe combined immunodeficiency[a] (SCID; *Chapters 4* and *5*)
Phenylketonuria (PKU; *Section 15.6*)
Sickle cell disease (*Chapter 13*)

[a] Only tested for in restricted areas of England.

Table 15.8 Genetic/inherited genetic clinical conditions tested for in the UK following birth.

15.10 CASE STUDIES

CASE STUDY 15.1

Christine was born after a normal pregnancy. For the first few days, she was healthy, but then started to vomit frequently. Her mother noted that her urine had a peculiar mousy smell. About a month following birth, Christine was admitted to hospital for the frequent vomiting, where it was realised that the genetic screening test after her birth had been deficient. Her blood was analysed for phenylalanine, with the result shown below (reference value in parentheses):

Plasma [phenylalanine]	1.6	(<0.1 mmol/L)

Questions

(a) What is the diagnosis for Christine?
(b) How should Christine be treated?
(c) Why does Christine's urine have a characteristic mousy odour?
(d) What would happen to Christine if she were not treated?

CASE STUDY 15.2

Jane is a 39-year-old healthy woman who has just given birth to her first child, a son, Peter. She had wanted the pregnancy in her words 'to be completely natural' and had refused an amniocentesis. Peter shows the very obvious physical characteristics of Down syndrome. *Figure 15.41* shows Peter's karyogram.

Questions

(a) Examine the karyogram in *Figure 15.41*. What is the cause of Down syndrome in this case?
(b) Given that a clinical diagnosis of Down syndrome is non-problematical, why is it necessary to obtain a karyotype?
(c) What advice would you give Jane?

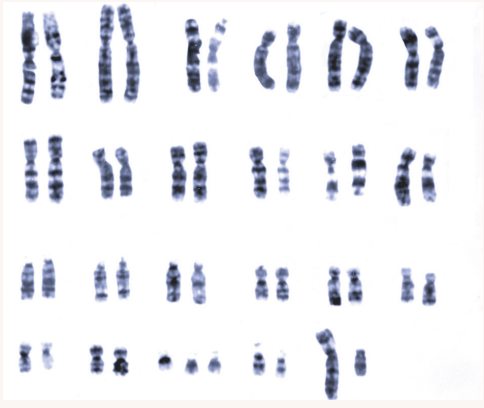


Figure 15.41 Peter's karyogram.

CASE STUDY 15.3

Jacob, a 20-year-old patient, visited his general practitioner (GP) after complaining of recurring pains in the lower back and joints. These symptoms, together with several physical signs, led his GP to refer him for a hospital consultation. His examination showed several significant diagnostic characteristics. For example, Jacob was above-average height but underweight for his age and gender, and had disproportionally long limbs. He displayed a positive Steinberg sign, that is, the tip of a thumb extended beyond the far edge of the palm when the hand was clasped. Jacob also showed a positive Walker–Murdoch wrist in being able to cover the nail of the little finger with his thumb when encircling his opposing forearm. He was short sighted and showed some dislocation of the lens (ectopic lens). An echocardiogram showed some cardiac structural abnormalities and a slight distention of the aorta in the lumbar region. This range of features suggested a diagnosis of Marfan syndrome (MFS). However, given there was no family history of MFS, a molecular biological test was requested, which quickly confirmed the diagnosis.

Marfan syndrome is a disorder of fibrous connective tissues. It is caused by mutations in *FBN1*/*MFS1* located at 15q21.1. This gene encodes the glycoprotein fibrillin, which is essential for the structural integrity and appropriate functioning of connective tissues. Mutations in *FBN1*/*MFS1* result in deficiencies of fibrillin or an unstable glycoprotein susceptible to proteolysis.

Questions

(a) What was the most likely molecular biology technique used to confirm the diagnosis?
(b) Is MFS a sex-linked condition?
(c) Would you expect it to be inherited in a Mendelian manner?
(d) Given that MFS is caused by mutations within a single gene, account for the diverse signs and symptoms associated with the condition.

15.11 SUMMARY

The sequence of nucleotides of the genes in the DNA forms the genetic blueprint of cells, controlling all their activities. This genetic information is transmitted to offspring in the gametes. Alterations in the sequence of nucleotides or to the chromosomal content of cells, that is, mutations, can lead to disease and/or death. Mutations may result in an absence or abnormality of a protein, perhaps an enzyme. Chromosomal aberrations involve structural changes to individual chromosomes, such as deletions, duplications, inversions and translocations. Changes to the chromosomal complement include aneuploidy and euploidy. Down syndrome is the most common aneuploid condition and is characterised by trisomy of the whole, or part of, chromosome 21. Genetic diseases can be diagnosed by karyotyping and by using molecular techniques, including the PCR, DNA hybridisation and next-generation sequencing methods. Once a genetic defect has been diagnosed, genetic counselling can be offered to affected individuals as part of the treatment process.

QUESTIONS

1. Which **ONE** of the following statements about PKU is **INCORRECT**?
(a) PKU sufferers should be placed on a diet low in phenylalanine.
(b) PKU has an autosomal dominant mode of inheritance.
(c) Individuals with untreated PKU may suffer from intellectual disabilities.
(d) PKU has an incidence of around 1:10 000 in the UK.
(e) PKU sufferers have insufficient melanin in their skin and hair.

2. Which of the following has an aneuploid karyotype?
(a) A male with a balanced Robertsonian translocation involving chromosomes 13 and 21.
(b) A male sufferer of PKU.
(c) A female sufferer of fatal familial insomnia.
(d) A female with a balanced translocation involving chromosome arms 11q and 22q.
(e) A male with Klinefelter syndrome.

3. Duchenne muscular dystrophy is an X-linked recessive trait. On average, what proportion of the children of a normal father and carrier mother would be affected?

4. A female is heterozygous for fragile X syndrome. Describe the inheritance pattern you would expect to see in her children if the father is a normal male.

5. Autosomal dominant and X-linked dominant diseases can both be passed on to children. Suggest characteristics that would help identify whether a particular syndrome was inherited in an autosomal or X-linked fashion.

6. Examine the three karyotypes ([A]–[C]) in *Figure 15.42*. In each case, name the syndrome caused by these karyotypes.

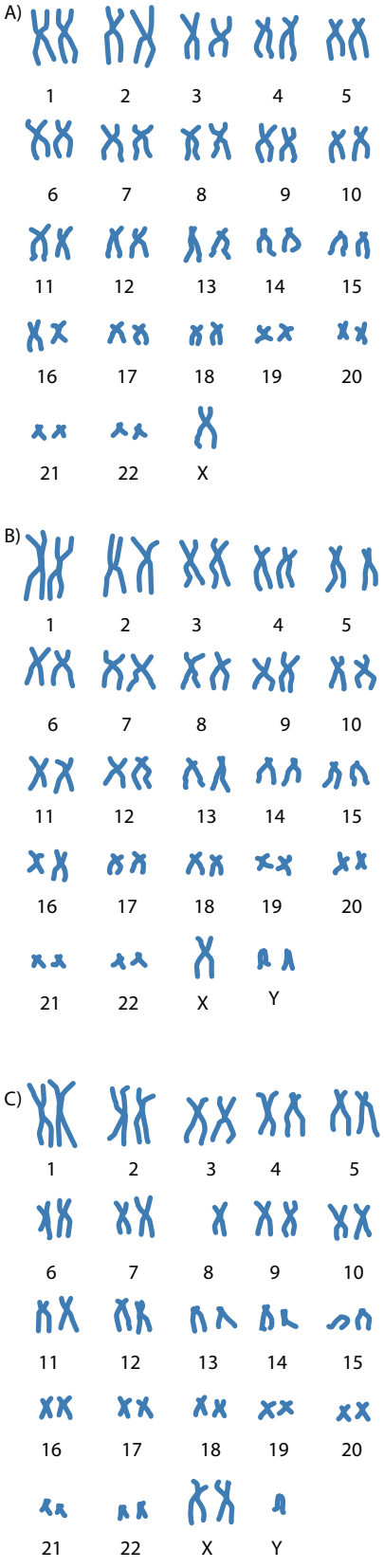


Figure 15.42 Karyotypes associated with question 6.

7. Arrange the following two lists into their most appropriate pairings.

Circular DNA molecules	Mitochondria
Somatic cell	Diploid
Synthesis of RNA and protein	S phase
RAD51	Chiasma
PKU	>0.7 mmol/L
PML	Nuclear bodies
Retinoid receptor	17q21.2
Nullisomy	2N–2
47,18+	Edwards syndrome
Down syndrome	Robertsonian translocation

FURTHER READING

Bull MJ (2020) Down syndrome. *N Engl J Med* **382**:2344–2352.

Cuckle H (1999) Maternal age-standardisation of prevalence of Down's syndrome. *Lancet* **354**:529–530.

Dabrowski E *et al*. (2019) Turner syndrome systematic review: spontaneous thelarche and menarche stratified by karyotype. *Horm Res Paediatr* **92**:143–149.

D'Aversa E *et al*. (2022) Droplet digital PCR for non-invasive prenatal detection of fetal single-gene point mutations in maternal plasma. *Int J Mol Sci* **23**:2819. doi: 10.3390/ijms23052819.

di Masi A *et al*. (2015) Retinoic acid receptors: from molecular mechanisms to cancer therapy. *Mol Aspects Med* **41**:1–115.

Gravholt CH *et al*. (2019) Turner syndrome: mechanisms and management. *Nat Rev Endocrinol* **15**:601–614.

Hanson B *et al*. (2022) Non-invasive prenatal diagnosis (NIPD): how analysis of cell-free DNA in maternal plasma has changed prenatal diagnosis for monogenic disorders. *Clin Sci (Lond)* **136**:1615–1629.

Ji F *et al*. (2022) New era of mapping and understanding common fragile sites: an updated review on origin of chromosome fragility. *Front Genet* **13**:906957. doi: 10.3389/fgene.2022.906957.

Kumar R *et al*. (2019) HumCFS: a database of fragile sites in human chromosomes. *BMC Genomics* **19**(suppl. 9):985. doi: 10.1186/s12864-018-5330-5.

Lee JM *et al*. (2021) Unravelling the physiological roles of retinoic acid receptor-related orphan receptor α. *Exp Mol Med* **53**:1278–1286.

Liquori A *et al*. (2020) Acute promyelocytic leukemia: a constellation of molecular events around a single PML-RARA fusion gene. *Cancers* (Basel) **12**:624. doi: 10.3390/cancers12030624.

Liverani MEC *et al*. (2019) Children and adults affected by Cri du Chat syndrome: Care's recommendations. *Pediatr Rep* **11**. doi: 10.4081/pr.2019.7839/.

Mackenzie G and Will R (2017) Creutzfeldt-Jakob disease: recent developments. *F1000Res* **6**:2053. doi: 10.12688/f1000research.12681.1.

National Genomic Test Directory (2023) Testing Criteria for Rare and Inherited Disease (V5). https://www.england.nhs.uk/publication/national-genomic-test-directories/.

Sigurdson CJ *et al*. (2019) Cellular and molecular mechanisms of prion disease. *Annu Rev Pathol* **14**:497–516.

Uttley LR *et al*. (2020) Creutzfeldt-Jakob disease: a systematic review of global incidence, prevalence, infectivity, and incubation. *Lancet Infect Dis* **20**:e2–e10.

van Spronsen FJ *et al*. (2021) Phenylketonuria. *Nat Rev Dis Primers* **7**:36. doi: 10.1038/s41572-021-00267-0.

Vardy ERLC *et al*. (2020) Phenylketonuria, co-morbidity, and ageing: a review. *J Inherit Metab Dis* **43**:167–178.

Wang H *et al*. (2019) Human prion disease. *Curr Opin Infect Dis* **32**:272–276.

Wang Q *et al*. (2022) Molecular biology techniques. In: Ahmed N *et al*. (eds), *Biomedical Science Practice*, 3rd edition (eds). Oxford University Press.

Useful websites

https://www.england.nhs.uk/publication/national-genomic-test-directories
https://www.cjd.ed.ac.uk
https://www.omim.org

Membrane and cell-related disorders

16

OBJECTIVES

After studying this chapter you should be able to:

- list some clinical conditions arising from specific defects in membranes and subcellular structures
- relate the symptoms of some disorders of membranes and subcellular structures to specific defects in their structures and activities
- discuss the diagnosis, management and treatment of some membrane and subcellular disorders

16.1 INTRODUCTION

Biological membranes are essential for life. The plasma membrane is the boundary that separates cells from their environment. The interior of eukaryotic cells is divided into structural and functional regions, the endomembrane system and organelles by membranes or double-membrane systems called envelopes. All biological membranes share a common basic structure but differ in their individual compositions, which are related to their functions. Indeed, genome sequencing projects have indicated that as many as 25% of all protein-coding genes may specify the structures of membrane proteins. Defects in their compositions or structures lead to clinical problems that in many cases are severe or even fatal. The shapes and locomotion of cells depend on a highly organised arrangement of fibrous proteins called the cytoskeleton. This is also responsible for the active transport of many materials around the cytoplasm. Some clinical conditions associated with the plasma and endomembranes, specific organelles and the cytoskeletal will be discussed.

16.2 CLINICAL DISORDERS ASSOCIATED WITH THE PLASMA MEMBRANE

Numerous clinical conditions are associated with defective receptors of the plasma membrane, the transport of materials across the membrane and/or subsequent dysfunctional signal transduction. Defects in receptor proteins and signal transduction can result in the development of cancerous states (*Chapter 17*), while cases of type 2 diabetes are associated with an ineffective insulin receptor (*Chapter 7*). Familial hypercholesterolaemia, described in *Chapter 14*, arises from impaired receptors for low-density lipoprotein (LDL) particles. Tangier disease, outlined in *Margin Note 14.5*, is caused by mutations in *TGD* sited at 9q31.1, which encodes a cholesterol transport protein of the plasma membrane. This protein is one of a group of transporters called ATP-binding cassette (ABC) proteins. These are multidomain structures: all have two cytosolic ATP-binding domains (NBDs [N, nucleotide]) or cassettes and two transmembrane domains, each consisting of several helices (*Figure 16.1*). One intensively studied ABC transporter is the cystic fibrosis transmembrane conductance regulator (CFTR), which functions as a Cl^- transporter and is gated by cycles of ATP binding and hydrolysis at its NBDs. Unlike Tangier disease, which is extremely rare, a dysfunctional CFTR leads to the much more common condition: cystic fibrosis (CF). This protein is a cAMP-regulated Cl^- channel whose structure has been determined (*Figure 16.2*).

DOI: 10.1201/9781003296836-16

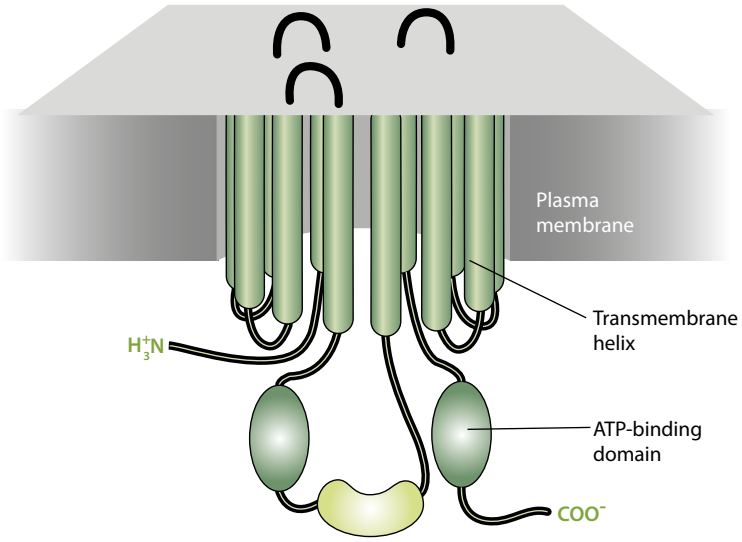


Figure 16.1 Schematic of an ATP-binding cassette protein. Redrawn from Akabas MH (2000) Cystic fibrosis transmembrane conductance regulator. *J Biol Chem* 275:3729–3732. Published under CC BY 4.0 licence. ⏎

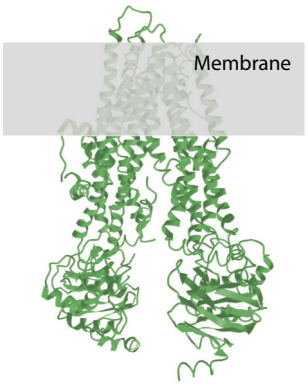


Figure 16.2 Molecular model of the cystic fibrosis transmembrane conductance regulator (CFTR). PDB file 5UAK. Compare with *Figure 16.1*. ⏎

Cystic fibrosis

Cystic fibrosis is a multisystem inherited disorder characterised by thick viscous mucus secretions in the lungs, digestive tract and associated organs, and epididymis. It is the most common fatal, homozygous recessive disorder of the white population, affecting about 1 in 2000–3000 people in the UK. The prevalence of CF varies throughout the world, with some ethnic groups reporting higher incidences. The onset may be at birth or later in childhood.

Cystic fibrosis results in chronic respiratory disease, malabsorption, cirrhosis and electrolyte disturbances (*Table 16.1*). It is characterised by decreased permeability of the apical membrane of epithelial tissues lining the lungs and other organs to Cl^- that results from mutations (*Chapter 15*) in *CFTR* sited at 7q31.2. Approximately 1 in 20 white people carry a copy of the mutated gene, and most of its 2000 mutations are known. The largest number of mutations are missense types. Indeed, the most common, which accounts for 70–90% of CF cases, is ΔF508, a deletion of a codon in exon 10 for an F (phenylalanine) residue at position 508 in the CFTR. The decreased permeability of membranes in CF patients results in many exocrine secretions with altered ion and reduced water contents and high viscosities. The appropriate viscosity is essential for normal functions of lubrication, digestion and protection by proteins secreted at epithelial surfaces. The highly viscous secretions can also obstruct ducts and lead to an accumulation of mucus and colonisation by bacteria.

Feature	Consequences
Recurrent respiratory infections	Irreversible lung disease
Pancreatic insufficiency	Malabsorption, poor growth, steatorrhoea (excessive fat in faeces)
Intestinal obstruction in neonates (meconium ileus)	Failure to thrive
Diabetes mellitus (late onset)	Failure to control blood sugar levels
Epididymal epithelial malfunction	Male sterility

Table 16.1 Some clinical features of cystic fibrosis. ⏎

Diagnosis and treatment of cystic fibrosis

Symptoms of CF usually become apparent soon after birth, and the majority of cases are diagnosed by the age of two years. Identifying mutations in the *CFTR* of a potential

BOX 16.1 CYSTIC FIBROSIS AND CHOLERA

Given that cystic fibrosis (CF) is essentially a fatal disease, one may wonder why mutated forms of *CFTR* persist in the population. One theory is that such mutations increase resistance to cholera (*Chapters 2* and *3*). The toxin of *Vibrio cholerae* (*Figure 16.3*) binds to surface receptors on the plasma membranes of gastrointestinal tract (GIT) enterocytes, leading to a voluminous secretion of fluid and Cl^-, resulting in a secretory diarrhoea that can be fatal if untreated. Cystic fibrosis mice that do not express *cftr* do not secrete fluid in response to cholera toxin. Heterozygous CF mice that express half the normal amount of cftr protein secrete about half the normal fluid and Cl^- in response to cholera toxin. Thus, CF heterozygotes might have a selective advantage in being resistant to cholera.

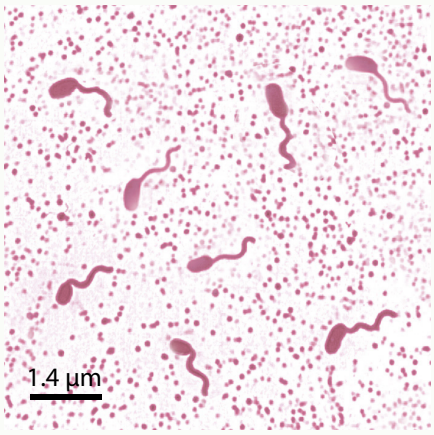


Figure 16.3 Photomicrograph of *Vibrio cholerae*. Courtesy of the Centers for Disease Control and Prevention.

CF patient will confirm the diagnosis and can also assist with carrier testing of the relatives of the patient. It is also used to give an accurate and rapid prenatal diagnosis of CF once the parental genotypes have been confirmed. In principle, any method capable of detecting a change in the DNA could be used to diagnose CF. Modern DNA sequencing methods (*Chapter 17*) are cheap and rapid. Methods based on the polymerase chain reaction (PCR; *Chapter 3*) are also effective. A simple way to detect the ΔF508 mutation is by amplifying that part of the gene and looking for differences in the sizes of the PCR products. Polyacrylamide gel electrophoresis (PAGE) can be used to separate the 98 bp product of PCR of the small fragment of exon 10 of the native *CFTR*, which may be visualised with ethidium bromide or silver staining. However, if the ΔF508 mutation is present, the PCR product is only 95 bp long. Careful PAGE can easily distinguish this relatively small difference (*Figure 16.4*). This test is not absolutely specific for ΔF508 given that other 3 bp deletions in this region would produce the same difference in size. Thus, the mutation ΔI507, the second identified CF mutation, also gives a 95 bp PCR product. Mutations other than ΔF508, however, can be identified in many cases because they produce differing banding patterns following PAGE.

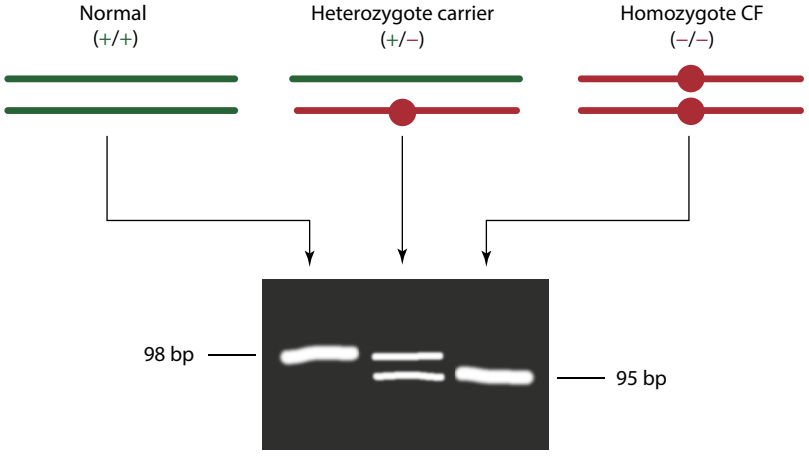


Figure 16.4 Detection of the ΔF508 mutation. Note the difference in sizes of the relevant portions of normal and mutated genes and how this allows their separation by polyacrylamide gel electrophoresis (PAGE).

When it is not possible to detect mutant *CFTR* alleles, then the sweat test remains the definitive test for diagnosis of CF. A diagnosis of CF can be confirmed by demonstrating a concentration of Cl^- in sweat greater than 70 mmol/L. Pilocarpine (*Figure 16.5*) can be

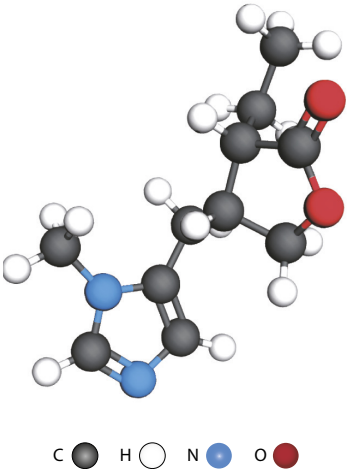


Figure 16.5 Computer drawn model of pilocarpine.

applied topically to stimulate the flow of sweat, samples of which are then collected and analysed for Cl^-. Unfortunately, the sweat test is not always reliable in the first six weeks of life or in adulthood. Furthermore, some patients with CF, including a number with severe lung disease, give a normal sweat test.

Several screening tests for CF are available based on the effects of the condition on pancreatic function. Measurements of immunoreactive trypsin (IRT) in dried blood spots, similar to those collected for phenylketonuria (PKU; *Chapter 15*) may also be confirmative. Immunoreactive trypsin is greatly increased in infants with CF in the first month of life compared with that of healthy infants. However, the test cannot be used after the first few weeks of life, since IRT levels fall as pancreatic insufficiency develops.

The general management of CF is aimed at preventing respiratory infections by regular percussive physiotherapy to try to expel excessive mucus, along with the use of antibiotics to counter bacterial infections and reduce inflammation (*Chapters 3* and *4*). Inhalation of hypertonic saline can be used to hydrate and reduce the viscosity of mucus in respiratory passages. Diets high in energy and salt, as well as vitamin supplements, are recommended. Pancreatic enzymes, in tablet form, may be given with food to counter the effects of pancreatic insufficiency. In recent years, low-M_r drugs called CFTR modulators have entered therapeutic use. These drugs target specific CFTR defects to improve their physiological performance; therefore, the type used is determined by the underlying mutation(s). Current modulators, such as amplifiers, correctors and potentiators, cannot fully restore Cl^- transport but do help to reduce the severity of symptoms. Amplifiers increase the amount of CFTR made by the cell. Correctors assist the CFTR in maintaining an appropriate active conformation, allowing it to function longer, while potentiators help Cl^- move more efficiently through the CFTR channel to the cell surface. The basis of treatment is to use modulators in appropriate combinations to improve the flow of Cl^- across the plasma membrane. Two correctors, lumacaftor and tezacaftor, are used in combination with the potentiator ivacaftor (*Figure 16.6*) to treat most CF patients over 12 years of age. The corrector elexacaftor is also used in combination with ivacaftor and tezacaftor. Previously, CF was usually fatal by the age of about 20 years, but these advances in therapy have produced marked improvements in life expectancy.

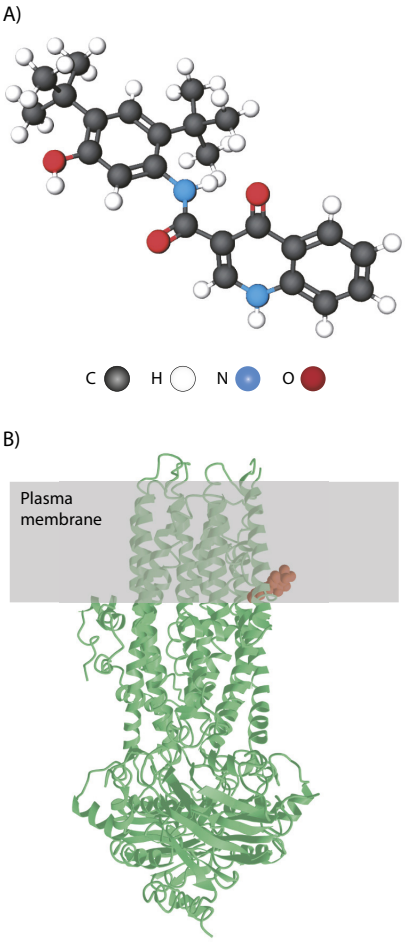


Figure 16.6 (A) Computer drawn model of ivacaftor and (B) molecular model showing its binding (highlighted in red) to the cystic fibrosis transmembrane conductance regulator (CFTR). PDB file 6O2P. Compare with *Figure 16.2*.

16.3 CLINICAL DISORDERS ASSOCIATED WITH THE ENDOMEMBRANE SYSTEM

The endomembrane system consists largely of the endoplasmic reticulum (ER), the Golgi apparatus (GA), lysosomes and a variety of vesicles. The envelope surrounding the nucleus (*Section 16.4*) is continuous with the ER and is sometimes considered a specialised part of its structure.

Disorders of the endoplasmic reticulum

The ER is a continuous system of single flattened membrane sacs and tubules (*Figure 16.7*). It also interacts with organelles, such as mitochondria, and other subcellular structures, for example the cytoskeleton, through ER proteins at membrane contact sites (MCSs).

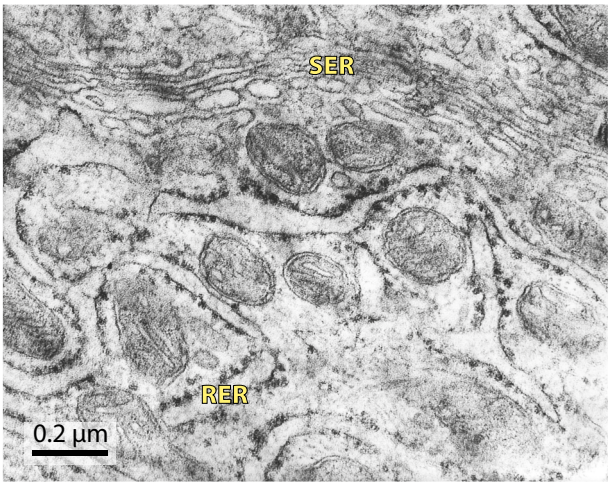


Figure 16.7 Electron micrograph showing a portion of the endomembrane system consisting of smooth and rough endoplasmic reticulum, (SER and RER, respectively). Note the numerous associated vesicles.

Margin Note 16.1
Deleted mice

Mice with deletions of *cftr*, that is, gene knockout mice, have been bred to be used as experimental models to test possible CF therapies. However, these mice do not develop lung disease in the same manner as humans and therefore do not provide an effective animal model. This lack has hampered experimental gene therapy studies that have tried to insert the wild-type gene into the lung epithelial cells of CF patients by using DNA-containing liposomes. To date, none of these therapies have been successful.

The mainly tubular smooth ER (SER) is adapted for the biosynthesis of lipids and accumulates and stores Ca^{2+}. Ribosomes of the rough ER (RER) synthesise polypeptides that are destined to become integral proteins of the plasma, ER, GA and lysosomal membranes. The interior or lumen of the RER is specialised for processing polypeptides to form proteins. It is an oxidising compartment that contrasts with the reducing environment of the cytosol. Within it, polypeptides are converted to native proteins by their appropriate folding and a variety of post-translational modifications including oxidation to form disulphide bonds and glycosylation to produce glycoproteins. The numerous activities of the ER are facilitated and catalysed by permeases and enzymes.

Attacks by parasites and pathogens (*Chapters 2* and *3*) and a variety of mutations can all impede ER activities. Viruses, such as some enteroviruses, and flaviviruses, like dengue and zika, bind to specific MCS proteins, which promotes viral replication. The replication of some bacteria, for example, *Chlamydia* and *Legionella pneumophila*, the causative organism of Legionnaires' disease, is also facilitated by their binding to MCS proteins. Mutations in genes for proteins that are necessary for appropriate ER structure and functions are implicated in numerous clinical conditions (*Table 16.2*). Additionally, ER stress occurs when the accumulation of unfolded polypeptides or misfolded proteins in its lumen exceeds a set threshold. This stimulates the unfolded protein response (UPR), an intracellular signalling pathway that adjusts the folding capacity of the ER and restores its homeostatic balance. However, high concentrations of misfolded proteins and chronic ER stress can induce cell death by **apoptosis** (*Chapter 17*) and are implicated in several chronic conditions, for example cardiovascular disease (*Chapter 13*), cancer (*Chapter 17*), diabetes (*Chapter 7*), fatty liver disease, inflammatory bowel disease (*Chapter 11*) and neurogenerative disorders (*Chapter 18*).

Clinical condition	Mutated gene (site)	Encoded protein	Molecular impairment
Amyotrophic lateral sclerosis	*RTN4* (2p16.1)	Replicon 4	Inappropriate distribution of protein disulphide isomerase
	STPB (9p13.3)	Sigma 1 receptor	Disrupts ER-mitochondrial MCSs leading to dysfunctional Ca^{2+} homeostasis
Spastic paraplegias	*ATL1* (14q22.1)	Atlastin 1	Disrupts formation of the tubular (ER) network
	KIF5A (12q13.3)	Neuron specific kinesin heavy chain	Impairment of axonal transport
	REEP1 (2p11.2)	Receptor expression enhancing protein 1	Disrupts ER-mitochondrial MCSs
	REEP2 (25q31.2)	Receptor expression enhancing protein 2	Inappropriate ER morphology
Charcot–Marie–Tooth disease (CMT) type 2	*MFN2* (1p36.22)	Mitofusin	Reduced ER-mitochondria MCSs
Stormorken syndrome	*STIM1* (11p15.4)	Stromal interaction molecule 1 ($[Ca^{2+}]$ sensor)	Dysfunctional intracellular Ca^{2+} metabolism
Tubular aggregate myopathy 1	*STIM1* (11p15.4)	Stromal interaction molecule 1 ($[Ca^{2+}]$ sensor)	Dysfunctional intracellular Ca^{2+} metabolism
Tubular aggregate myopathy 2	*ORAI1* (12q.24.31)	Ca^{2+} release activated Ca^{2+} modulator 1	Dysfunctional intracellular Ca^{2+} metabolism

MCS, membrane contact site.

Table 16.2 Examples of clinical conditions associated with mutated endoplasmic reticulum (ER) genes. ⏎

Disorders of the Golgi apparatus

The GA consists of flattened sacs of smooth membrane called cisternae that are arranged in stacks termed dictyosomes (*Figure 16.8*). The stacks are surrounded by vesicles and interconnected by tubules. The structure of the GA is maintained by its surrounding matrix, which contains so-called Golgi reassembly and stacking proteins (GRASPs) and by interactions with microtubules of the cytoskeleton (*Section 16.7*). The GA is the major site for sorting and distributing lipids and proteins. It receives material from the ER in transition vesicles, a process referred to as anterograde transport. Within the GA, biological materials are processed and sorted into a variety of vesicles for distribution to

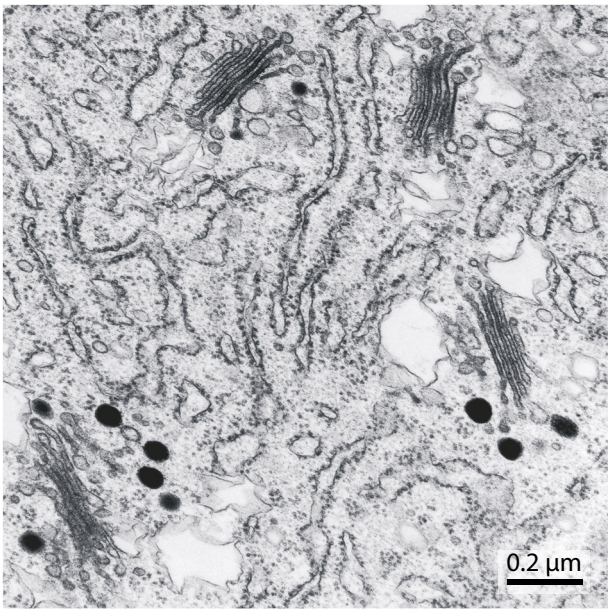


Figure 16.8 Electron micrograph showing four Golgi apparatus cisternae.

other intracellular sites, including the ER (retrograde transport), to endosomes to form the lysosomal system and to the plasma membrane. These diverse transport activities are usually referred to as trafficking. A family of proteins called golgins are bound to the GA membrane at their carboxy termini and form a complex network surrounding the GA. These large proteins have M_r greater than 500 000 and adopt a mainly coiled-coil conformation (*Margin Note 16.2*) forming molecular rods that can extend up to 400 nm from the Golgi membrane. Golgins have binding sites along their lengths, allowing them to recognise and specifically bind different types of vesicles, and so are essential to the role of the GA in vesicle trafficking. In addition to trafficking, the GA has several biosynthetic roles, especially in the post-translational modification of proteins. For example, it contains glycosyltransferases and proteases. The former complete the formation of glycoproteins and glycolipids, while the latter catalyse the hydrolyses necessary for some inactive proteins to adopt active conformations, such as the conversion of proinsulin to insulin (*Chapter 7*).

Like the ER membrane, that of the GA contains several permeases that allow ions and substrates to enter its lumen. Also, like the ER, some clinical disorders are associated with dysfunctional GA proteins. For example, mutations in genes encoding several glycosyltransferases cause some types of muscular dystrophies (*Section 16.7*). Examples of other GA-related disorders are listed in *Table 16.3*. Defective GA glycosylation is apparent in many cancers. Apoptosis can be initiated by, for example, mutations in several golgin genes and contributes to the pathology of some chronic inflammatory diseases and neurodegenerative disorders. Membrane-bound transcription factor protease site 1 (MBTPS1), formerly called site-1 protease (S1P), is a serine protease in the GA encoded by *MBTPS1* (16q23.3-q24.1). Mutations that cause a deficiency in the activity of this enzyme lead to defective trafficking between the GA and ER and dysfunctional lysosome biogenesis with secretion of lysosomal enzymes into the plasma. Indeed, the best characterised clinical disorders of the endomembrane system are lysosomal storage diseases (LSDs).

Lysosomal storage diseases

Lysosomes are organelles limited by a single, impermeable membrane. They contain about 60 different hydrolases that are active at acid pH and can degrade all biological materials (*Figure 16.10*). Other membrane-bound vesicles, endosomes and phagosomes, fuse with lysosomes and deliver materials taken up by the cell by endocytosis or phagocytosis for degradation by these enzymes. Lysosomal enzymes are synthesised by the ribosomes of the RER and enzymes within the GA. In the GA, biosynthetic enzymes recognise sites on the surface of the prolysosomal enzymes and tag them with mannose 6-phosphate residues that act as molecular addresses, ensuring that they are integrated into vesicles that eventually become lysosomes.

Margin Mote 16.2
Coiled coils

Coiled coils consist of two to seven α helices tightly twisted together (*Figure 16.9*). The individual helices are held together largely by hydrophobic interaction and van der Waals forces. The existence of coiled-coil structural motifs in proteins was predicted on theoretical grounds in the early 1950s by Francis Crick (1916–2004), of DNA double-helical fame, long before the molecular structure of any protein was known.

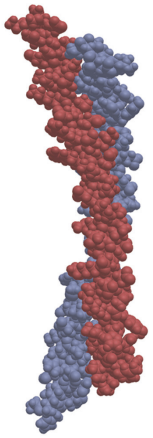


Figure 16.9 Molecular model of a coiled-coil motif. PDB file 2N9B.

Clinical condition	Mutated gene (site)	Encoded protein	Molecular/cellular impairment
Achondrogenesis, type IA; odontochondrodysplasia 1	*TRIP11* (14q32.12)	Golgin GMAP-210	GA fragmentation
Congenital disorder of glycosylation	*SLC35A2* (Xp11.23)	UDP Gal permease	Defective glycosylation
COPA syndrome	*COPA* (1q23.2)	COPA	Defective trafficking
Cutis laxa type IIC	*ATP6V1E1* (22q11,21)	Ion pump	Impaired GA–ER retrograde transport
MAC syndrome	*RIN2* (20p11.23)	Ras and Rab interactor 2	Abnormal GA morphology and trafficking causing collagen and elastin disorders
Menkes disease (*Chapter 10*); occipital horn syndrome	*ATP7A* (Xq21.1)	Cu^{2+}-ATPase transporter	Defective GA Cu^{2+} metabolism
Smith–McCort dysplasia 2	*RAB33B* (4q31.1)	RAB33B[a]	Defective trafficking and GA fragmentation

COPA, coatomer protein complex, α subunit; ER, endoplasmic reticulum; GA, Golgi apparatus; MAC, microphthalmia, anophthalmia and coloboma. [a] Ras-associated protein.

Table 16.3 Examples of clinical conditions associated with mutated Golgi apparatus (GA) genes. ⏎

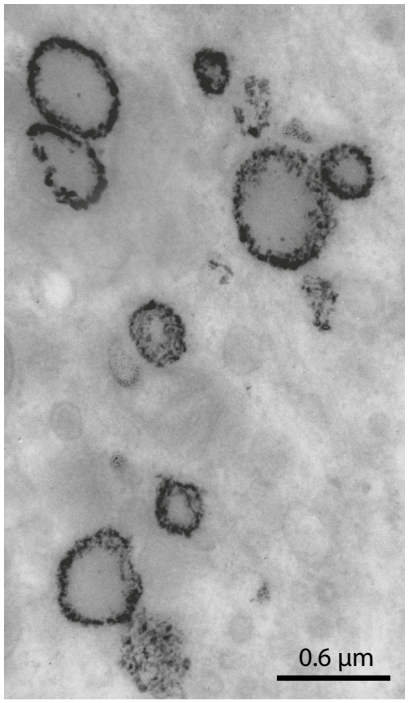


Figure 16.10 Electron micrograph of lysosomes in a synovial cell identified by aryl sulphatase activity. Courtesy of Dr C.J.R. Jones, Department of Pathology, University of Manchester, UK. ⏎

Lysosomal storage diseases are caused by mutations in genes encoding lysosomal hydrolases or, less often, in transporters that normally function in the export of enzyme products from lysosomes. In either case, the accumulation of substrate results in lysosome dysfunction and eventually cell death, which impairs the functions of major organs such as the brain and liver. In most cases, LSDs are caused by the deficiency of a single enzyme. They can be subdivided according to which pathway is affected. For example, enzyme deficiencies in the pathway for degrading glycosaminoglycans cause mucopolysaccharidoses and deficiencies affecting glycopeptides lead to glycoproteinoses; glycolipid storage diseases result from deficiencies in enzymes involved in the degradation of sphingolipids. More than 40 LSDs are known (*Table 16.4*). All are relatively uncommon

Disease	Enzyme deficiency	Material (substrate) accumulated
Farber granulomatosis	Ceramidase	Ceramides
Fucosidosis	α Fucosidase	Oligosaccharides, glycolipids
Gaucher disease	β Glucocerebrosidase	Glucosylceramide
Inclusion cell disease (I cell disease or mucolipidosis II)	Numerous lysosomal enzymes absent	Glycolipids, glycoproteins, sialyloligosaccharides
Metachromatic leukodystrophy	Arylsulphatase A	Sulphatides
Niemann–Pick disease	Sphingomyelinase	Sphingomyelin
Pompe disease type II	α Glucosidase	Glycogen
Sphingolipidosis GM_1 gangliosidosis	β Galactosidase	GM_1 gangliosides
Tay–Sachs disease (GM_2)	Hexosaminidase A	GM_2 gangliosides

Table 16.4 Examples of lysosomal storage diseases (LSDs). ⏎

in the general population, although several are more prevalent among Ashkenazi Jews. These disorders are congenital and seriously reduce the quality of life and life expectancy. For example, Tay–Sachs disease (*Table 16.4* and *Figure 16.11*) leads to blindness, paralysis, dementia and death, usually by the age of two years. However, although the substrates that accumulate differ between individual LSDs, in many cases the resulting clinical presentation and pathological features are similar. Gaucher disease (GD) shows many of these generic features.

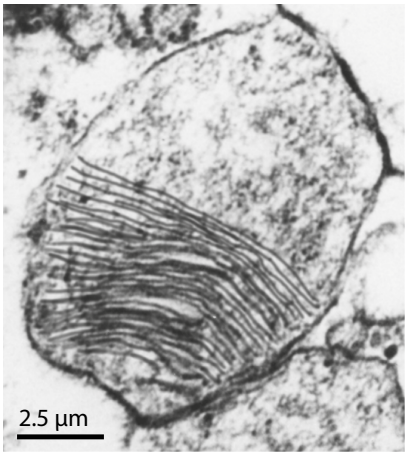


Figure 16.11 Electron micrograph of an enlarged lysosome containing a zebra body of accumulated lipid (GM_2 ganglioside; see *Table 16.4*) from the cerebral cortex of a 19-week-old fetus. A diagnosis of Tay–Sachs disease was made by showing a deficiency of hexosaminidase A activity in cultured amniotic fluid cells obtained by amniocentesis (see *Figure 15.40*), following which the fetus was aborted. Courtesy of Dr A. Cooper, Willink Biochemical Genetics Unit, Royal Manchester Children's Hospital, UK.

Gaucher disease

Gaucher disease, named after the French dermatologist Philippe Gaucher (1854–1918), is characterised by the presence of Gaucher cells (*Figure 16.12*), which are histiocytes of macrophage or Langerhans cell lineages (*Chapter 4*) enlarged with accumulated lipid. Gaucher disease is the most common LSD, with an incidence of 1 in 40 000–60 000, although it is about 50 times more common in Ashkenazi Jews. It is an autosomal inherited recessive condition caused by mutations in *GD* located at 1q22, which encodes a lysosomal β glucocerebrosidase. More than 300 mutations of this gene have been described. Deficiencies in the activity of β glucocerebrosidase lead to an accumulation of its substrate, glucosylceramide, within lysosomes and the formation of Gaucher cells in bone marrow, the reticuloendothelial system, parts of the circulation and other organs. The pathogenic mechanisms are poorly understood, but three forms are recognised: type 1 or adult (GD1), type 2 or infantile (GD2) and type 3 or juvenile (GD3), with the differences arising from different mutations in *GD*.

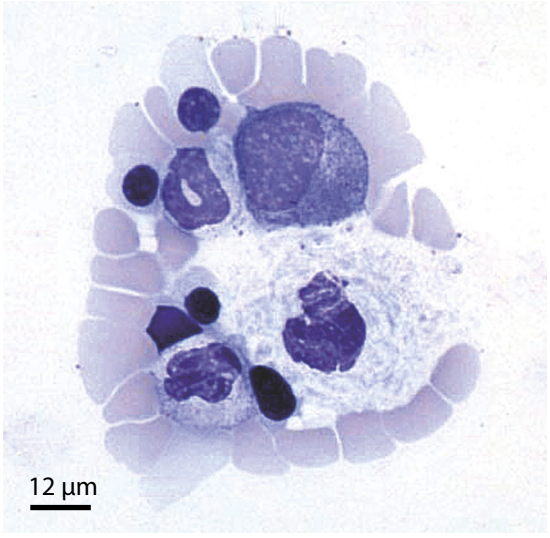


Figure 16.12 Photomicrograph of a Gaucher cell in a bone marrow aspirate. Note the abnormal cytoplasm containing heavily stained vesicles (see *Table 16.4*). Courtesy of Glenn Anderson, Great Ormond Street Hospital for Children, London, UK.

Type 1 GD is the most prevalent form. It may present from infancy to adulthood. Although some sufferers may die at a young age due to pulmonary infections, most survive to late adult life. The condition is characterised by hepatosplenomegaly, the enlargement of the liver typically occurring after that of the spleen. The abdomen can also enlarge due to distention of the colon. Most of these visceral features arise from macrophage dysfunction. The skin bruises very easily and becomes pale yellow in colour. Febrile episodes with generalised pain in the body and limbs occur. Painful bones and joints, degeneration of vertebral bodies and arthritic hip joints are common. Pathological fractures may also occur. Mild anaemia and recurrent pulmonary infections are characteristic. Some patients show neurological abnormalities and decreased intelligence, with ataxia and fits. Type 1 GD is associated with a higher risk of Parkinson disease, some solid cancers and haematologic diseases, particularly multiple myeloma (see *Figure 1.18* and *Chapter 17*). Type 2 GD is much more severe, with death occurring by the age of two years following progressive psychomotor regression. The earliest symptoms include a protuberant abdomen usually due to hepatosplenomegaly and a general failure to thrive. Indeed, after six months, developmental arrest and neurological regression are obvious characteristics. Fits may occur. The baby/infant becomes spastic, with laryngeal spasms causing difficulties in swallowing, and **strabismus** (a squint or abnormal alignment of the eyes). Deterioration is rapid, leading to recurrent pulmonary infections and death. Type 3 GD is the rarest of the three. Although it presents later than type 2 and has a slower progression, it also has a characteristic neurological deterioration. The major features are hepatosplenomegaly, convulsions and psychomotor retardation. Ataxia, spasticity and strabismus may also occur.

Diagnosis and treatment of lysosomal storage disorders

Early diagnosis is important to prevent serious damage to the nervous system and major organs. Relatively simple laboratory methods can accurately diagnose patients exhibiting the typical traits of LSDs. However, atypical patients usually require more detailed investigations. Almost all LSDs can be diagnosed using samples of leukocytes or plasma using assays to determine reduced activities of lysosomal enzymes or an increase in substrate in skin fibroblasts, plasma or urine. Prenatal diagnosis is possible when the likelihood of an LSD being present in a fetus is high. If this is the case, genetic counselling (*Chapter 15*) has a role, and a therapeutic abortion can be offered. Treatment of LSDs is supportive but, in several cases, enzyme replacement therapy (ERT) with or without substrate reduction therapy (SRT) has proved beneficial.

Gaucher disease may be diagnosed by demonstrating β glucocerebrosidase activity in leukocytes of only about 10% of reference values. This may be confirmed by identifying biallelic mutations in *GBA1*. The prognosis for adult GD patients is relatively good. Supportive treatment includes analgesics for the bone pains, blood transfusions to relieve the anaemia and orthopaedic relief for fractures and degeneration of the hip joints. First-line clinical treatment is ERT by the intravenous administration of recombinant β glucocerebrosidase (*Figure 16.13*). The enzyme is partially deglycosylated to expose mannose residues that are recognised by receptors on the surfaces of macrophages, which allows absorption of the enzyme into the cells. Within the macrophages, the enzyme is delivered to the lysosomes where it catalyses the breakdown of the accumulated glucosylceramide. In most cases, this is an effective and safe treatment, reducing the sizes of the liver and spleen and allowing them and the bone marrow to function effectively. Substrate reduction therapy uses molecules of low M_r that, as the name implies, decrease the production of substrate. For example, miglustat and eliglustat (*Figure 16.14*) are used to treat GD because they inhibit glucosylceramide synthase, which prevents the accumulation of glucosylceramide in lysosomes. The general management of GD patients may involve splenectomy to relieve the hypersplenism, although this may hasten bone deterioration. Bone marrow transplants (*Chapter 6*) are not generally considered because of the effectiveness of ERT and SRT. Gene therapy, in which a functional *GD* is inserted into stem cells in the bone marrow, has, to date, not proved effective.

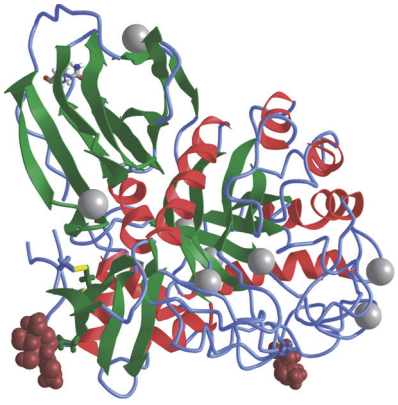


Figure 16.13 Molecular model of a recombinant β glucocerebrosidase. Bound glycans are shown in brown; the grey spheres represent Mg^{2+}. PDB file 6TN1.

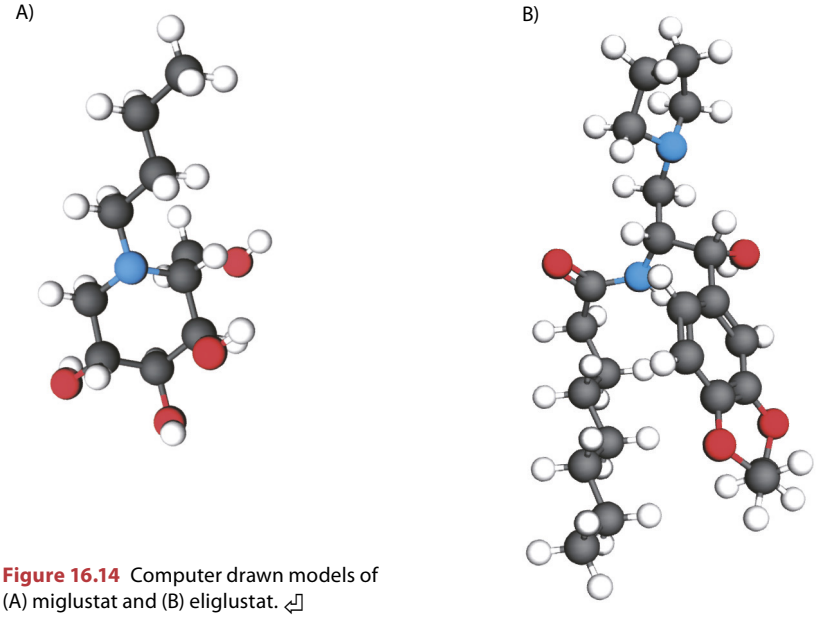


Figure 16.14 Computer drawn models of (A) miglustat and (B) eliglustat.

16.4 NUCLEI AND NUCLEOCYTOPLASMIC TRANSPORT

All eukaryotic cells possess a nucleus that contains its chromosomes. The nucleus is separated from the cytoplasm by a nuclear envelope consisting of outer and inner nuclear membranes that enclose a perinuclear space. The two membranes constitute a single system but have different compositions and functions. The inner membrane has receptors that bind a system of fibrous proteins called the nuclear lamina, which forms part of

the nuclear skeleton (*Section 16.7*). The outer membrane is continuous with the ER and, like its rough portion, is studded with ribosomes. The perinuclear space and the lumen of the ER form a single enclosed space (*Figure 16.15*).

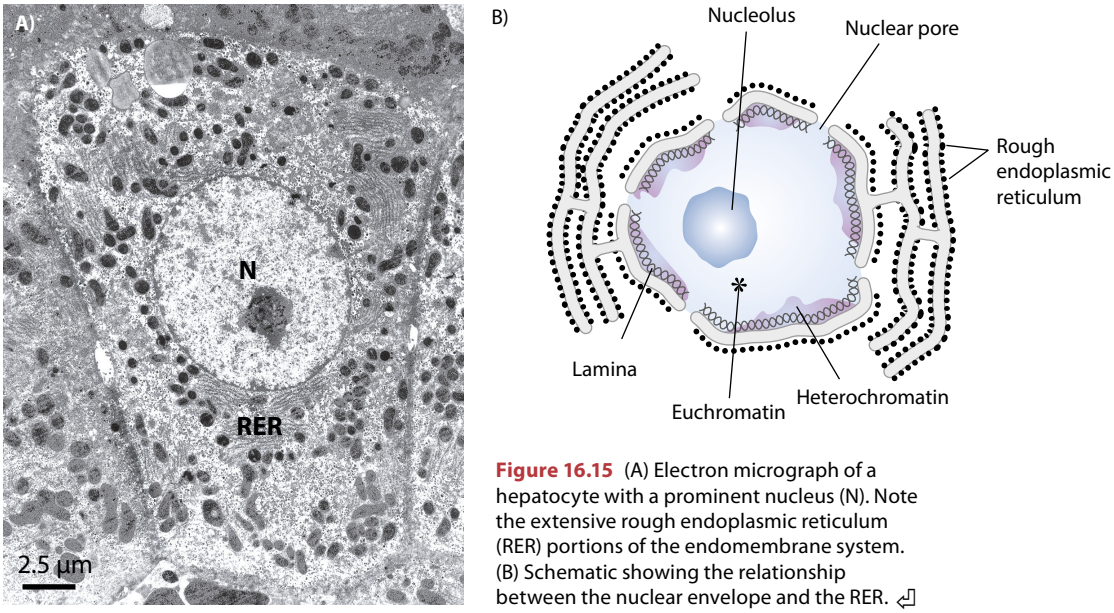


Figure 16.15 (A) Electron micrograph of a hepatocyte with a prominent nucleus (N). Note the extensive rough endoplasmic reticulum (RER) portions of the endomembrane system. (B) Schematic showing the relationship between the nuclear envelope and the RER.

Polypeptides synthesised by ribosomes on nuclei and ER surfaces are transferred into the periplasmic space and ER lumen, respectively.

Nuclear pores (NPs) in the envelope (*Figure 16.16*) allow the transport of materials between the nucleoplasm and cytoplasm, a process called nucleocytoplasmic transport. Thus, proteins, such as histones, and enzymes and proteins involved in the replication and transcription of DNA and those involved in processing initial RNA transcripts and the assembly ribonuclear particles can enter the nucleus. Conversely, ribosomal subunits and tRNA and mRNA molecules are exported from the nucleus to the cytoplasm.

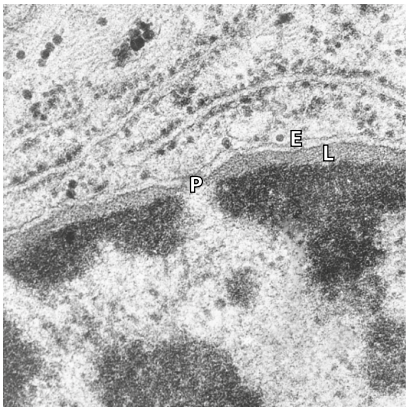


Figure 16.16 Electron micrograph showing the nuclear envelope (E), lamina (L) and a nuclear pore (P). Picture by G. Szabo from Fawcett DW (1981) *The Cell*, 2nd edition. WB Saunders, with permission. Courtesy of Emeritus Professor D.W. Fawcett.

Nuclear pores

Among the best-studied NPs are those from amphibians, such as the toad *Xenopus laevis*. Each pore has an eightfold symmetry. It consists of an NP complex (NPC) from which fibrils, 50–100 nm long, extend into the cytoplasm while a basket-like attachment extends into the nucleoplasm (*Figure 16.17*). An NPC consists of multiple copies of about 30 proteins called nucleoporins and has M_r of approximately 125×10^6. Complexes are cylindrical and consist essentially of cytoplasmic and nucleoplasm rings of nucleoporins either side of an inner ring, which form a channel of approximately 40 nm diameter through the envelope (*Figure 16.18*). Channel nucleoporins have numerous repeating pairs of FC (Phe-Gly) residues. These cannot form a regular secondary structure but consist of large extended regions forming a gel-like mesh that mediates transport through the pore. The 5000-or-so pores of a typical mammalian nucleus are held in place by attachments of nucleoporins to proteins in the nuclear envelope and to the nuclear lamina.

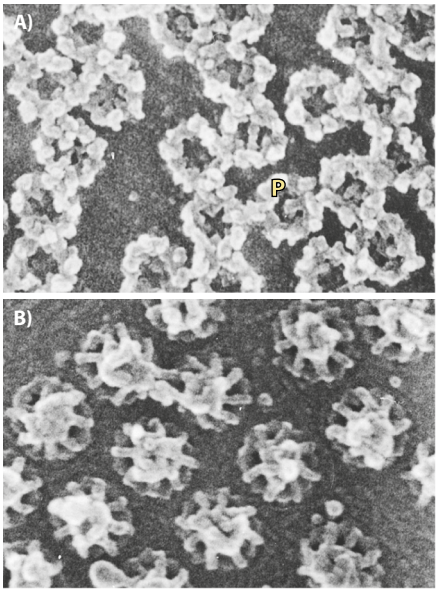


Figure 16.17 Electron micrographs of the (A) outer and (B) inner faces of nuclear pore (P) complexes approximately 90 nm in diameter. Courtesy of Drs T.A. Allen and M.W. Goldberg, Department of Structural Cell Biology, Paterson Institute, Manchester, UK.

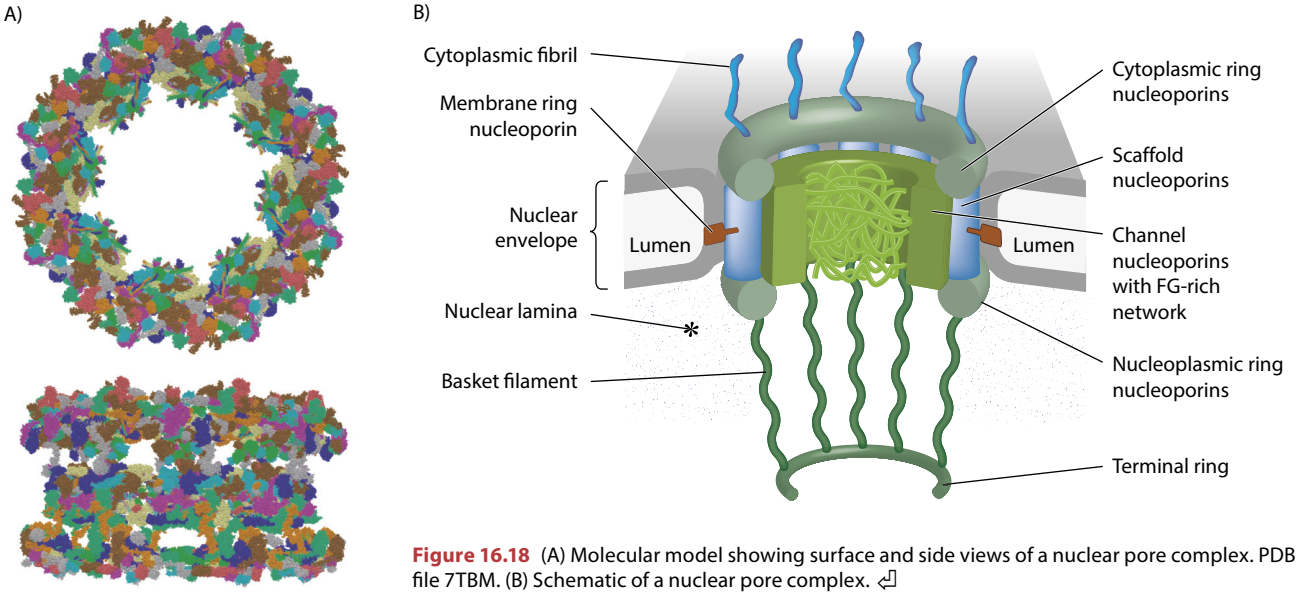


Figure 16.18 (A) Molecular model showing surface and side views of a nuclear pore complex. PDB file 7TBM. (B) Schematic of a nuclear pore complex.

Nucleocytoplasmic transport

Nucleocytoplasmic transport is a rather complicated process and only a simplified view is given here. The mesh occupying the NPC forms a gated channel. Molecules of less than M_r 5000 can pass freely through the channel and cross the envelope by diffusion. Generally, however, proteins with M_r larger than 30 000 cannot cross by free diffusion but they can enter the nucleus by active transport if they possess a nuclear locating signal (NLS). The best known NLSs are rich in basic amino acid residues, such as that of the large T antigen of the simian virus 40, which has the sequence ...P-K-K-K-R-K-V... (...Pro-Lys-Lys-Lys-Arg-Lys-Val...). A protein with a NLS is called a cargo. Nucleocytoplasmic transport also requires the participation of a number of soluble import factors, for example nuclear import receptors, which are proteins that recognise the NLS of cargos (*Figure 16.19*) and bind to them. Receptors also possess several binding sites that transiently interact with FC regions of channel nucleoporins and so facilitate passage of the complex through the NPC channel (*Figure 16.20*). Following transport through an NPC into the nucleus, the importin–cargo complex binds to a small G protein called Ran (*Chapter 7*). The GTPase activity of Ran allows it to bind and then hydrolyse a GTP to GDP. The conformation of Ran depends upon whether it has a bound GTP or a GDP.

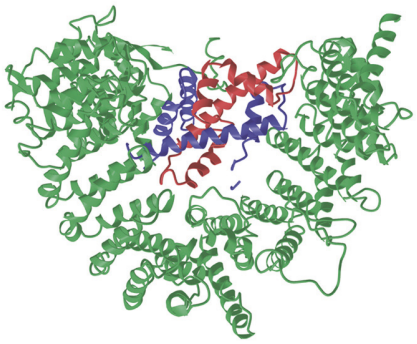


Figure 16.19 Molecular model of a nuclear import receptor, shown in green, with its cargo of two bound histones shown in red and blue. PDB file 6N1Z.

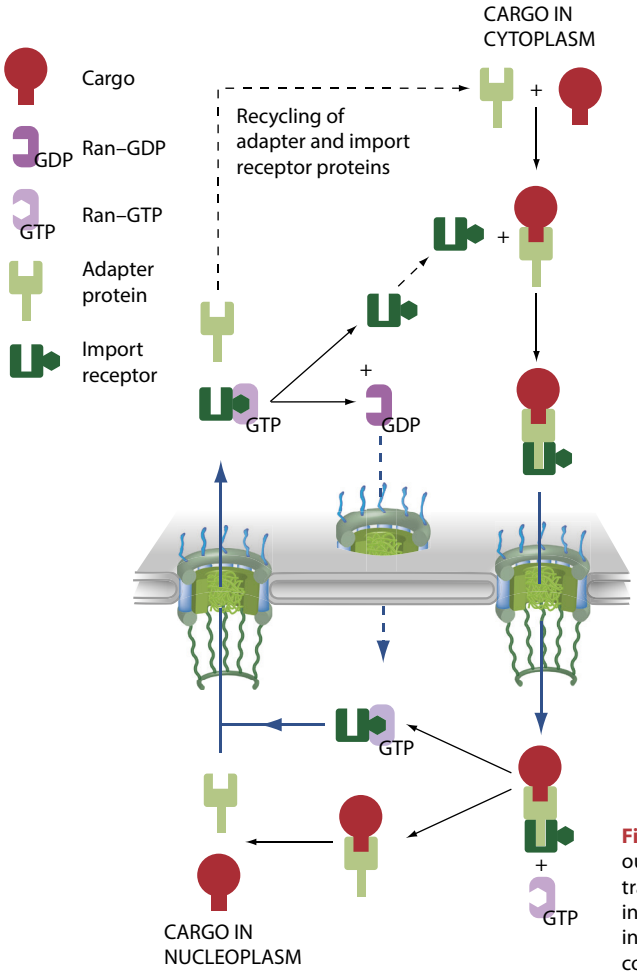


Figure 16.20 A simplified outline of the nucleocytoplasmic transport of a protein cargo into the nucleus. Blue arrows indicate the movements of the complexes.

The binding of Ran–GTP to the complex releases the cargo in the nucleus and allows the import receptor–Ran–GTP to be shuttled to the cytoplasm. Here, yet another protein, GTPase-accelerating/activating protein (Ran–GAP, not shown in *Figure 16.20* for the sake of clarity), stimulates Ran to hydrolyse its bound GTP to GDP. The conformation of Ran–GDP has only a low affinity for the import receptor, which is then released and can participate in a fresh cycle of nuclear import. The Ran–GDP is returned to the nucleus where it releases its GDP and binds a GTP under the influence of a specific guanine nucleotide-exchange factor (Ran–GEF).

The export of materials from nucleus to cytoplasm resembles import processes but in reverse. Nuclear export signals on cargos in the nucleus are recognised by nuclear *export* receptors, which bind to the cargo in a Ran–GTP-mediated process. Transport through the NPC is then facilitated. Within the cytoplasm, Ran–GAP stimulates the GTPase activity of Ran, converting it to its GDP-bound conformation. This has limited affinity for the exportin–cargo complexes, which are released, allowing the exportin to be returned, via a NPC, to the nucleus.

Note that the transport to and release of cargos in their respective destinations is dependent on the free energy of hydrolysis of Ran–GTP. The direction of transport into or out of the nucleus is driven by the concentration gradient of Ran–GTP across the envelope, being highest in the nucleoplasm and lowest in the cytoplasm. Nuclear pores are busy places. Hundreds, if not 1000s, of cargoes are moved through a single NPC/minute, and in both directions simultaneously.

Clinical aspects of nucleocytoplasmic transport

Malfunctioning nucleocytoplasmic transport is associated with several clinical conditions including Alzheimer and Huntington diseases (*Chapters 15* and *18*), where the expression of nucleoporins is dysfunctional following mutations in their genes or because they are inappropriately located. For example, in Huntington disease, a single gene (*HD* at 4p16.3) is mutated and the corresponding dysfunctional protein, huntingtin (*Figure 16.21*), accumulates in the nucleus due to the inappropriate activities of NPCs.

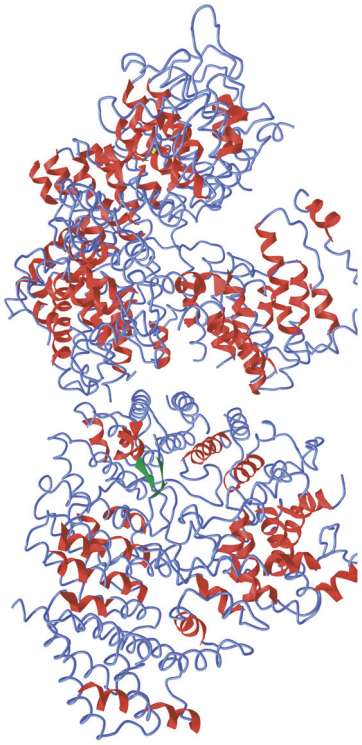


Figure 16.21 Molecular model of huntingtin. PDB file 6YEJ.

Nuclear pore complexes have also been implicated in several autoimmune diseases caused by autoantibodies against nucleoporins. This contributes to a number of clinical conditions, such as systemic lupus erythematosus, rheumatoid arthritis and primary biliary cirrhosis (*Chapter 5*).

Given that NPCs are the pivotal junctions between the nucleus and cytoplasm, they are closely involved in the pathology of viral infections. Virus particles fuse with the plasma membrane or are taken up into the cells by endocytosis (*Chapter 2*). However, they or their nucleic acid must leave the cytoplasm and enter the nucleus for viral replication to occur. For example, the hepatitis B virus (HBV) has a capsid of diameter 32–36 nm, which is small enough to move through an NPC directly. *In vivo*, phosphorylation of HBV capsids makes their NLSs accessible to the nuclear transport machinery. Following their transport through an NPC, capsid protein and DNA from mature viruses are released into the nucleoplasm. In contrast, adenovirus particles and herpes simplex virus (HSV) have capsids with diameters of 90 and 125 nm, respectively, which are too large to traverse an NPC. These virus particles dock at the cytoplasmic side of the NPC in an importin-dependent manner. The HSV then releases its DNA through the NPC into the nucleoplasm. Adenovirus particles, however, first trap a variety of transport factors that are necessary for the subsequent disassembly of the capsid before translocation of the viral DNA into the nucleus is possible.

Nucleoporins have been implicated in several types of cancers (*Chapter 17*). In some cases, these are associated with the overexpression of NPC proteins, but in most other cases tumorigenesis stems from chromosomal rearrangements that result in oncogenic fusion proteins. Examples include tumours, such as acute myeloid leukaemia, chronic myeloid leukaemia, inflammatory myofibroblastic tumour and myeloblastic leukaemia, together with adenocarcinomas of the breast, lung and pancreas, osteosarcomas and papillary thyroid carcinomas that affect a variety of tissues. Chromosomal translocations of the gene for nucleoporin 98 (*NUP98*) are frequently found in acute myeloid leukaemia and are associated with poor patient outcomes.

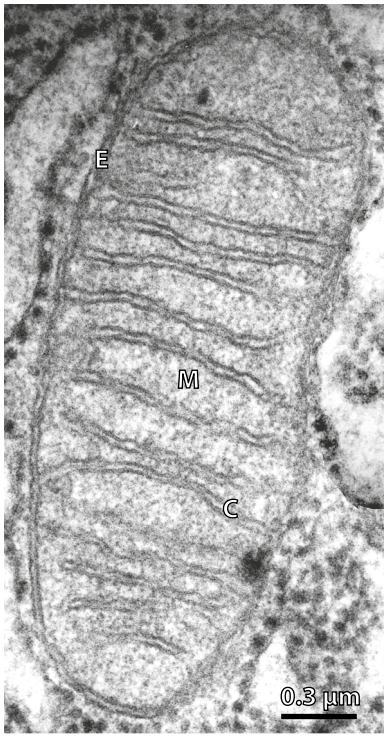


Figure 16.22 Electron micrograph of a mitochondrion. C, cristae; E, envelope; M, matrix. ⏎

16.5 MITOCHONDRIAL DISORDERS

Mitochondria are organelles found in almost all eukaryotic cells, although numbers vary from one to several hundred per cell. They are bounded by a mitochondrial envelope. This consists of outer and inner mitochondrial membranes. The latter has infoldings called cristae and encloses an inner compartment called the mitochondrial matrix (*Figure 16.22*). The inner mitochondrial membrane contains five protein complexes (*Table 16.5*) that together with cytochrome *c* perform energy transduction reactions with the formation of ATP by oxidative phosphorylation.

Complex	Name	Comment
Complex I	Reduced nicotinamide adenine dinucleotide (NADH): ubiquinone reductase	Complexes I and II transfer electrons from NADH and succinate, respectively, to the mobile carrier ubiquinone, which then transfers them to complex III.
Complex II	Succinate: ubiquinone oxidoreductase	
Complex III	Ubiquinone: ferricytochrome *c* oxidoreductase	Complex III transfers electrons to the second mobile carrier, cytochrome *c*, which transfers them to complex IV.
Complex IV	Cytochrome *c* oxidase	Complex IV donates electrons to dioxygen.
Complex V	ATP synthase	Complex V synthesises ATP from ADP and P_i when stimulated by H^+.

Table 16.5 The protein complexes of the inner mitochondrial membrane involved in oxidative phosphorylation. ⏎

The matrix contains enzymes that catabolise fuel molecules to yield the reduced coenzymes NADH and $FADH_2$ necessary for oxidative phosphorylation, and produces small organic molecules that are precursors in biosynthetic metabolism. The mitochondria also contribute to the intracellular homeostasis of many metabolites and ions, including Ca^{2+} and H^+. Mitochondria are also involved in cellular responses to environmental toxins that can stimulate them to release a potent mixture of enzymes and other proteins that eventually bring about the death of the cell by a form of apoptosis or cell

suicide (*Chapter 17*). These apoptotic mechanisms are, of course, under strict regulation, but defects in these controls contribute to some neurodegenerative and autoimmune diseases and cancers.

Mitochondrial DNA and its replication

Mitochondria are unusual organelles in that their matrix contains multiple copies of molecules of mitochondrial DNA (mtDNA). Human mtDNA is a circular double-stranded molecule that has been completely sequenced. It consists of 16 569 base pairs but contains only 37 mitochondrial genes (*Figure 16.23*) that are asymmetrically distributed: 28 encoded on one strand and nine on the other. Together, these genes encode 13 mitochondrial polypeptides that are involved in oxidative phosphorylation, 22 mitochondrial tRNAs, and the 12S and 16S rRNAs that form part of the ribosomes found in mitochondria. Mitochondrial ribosomes differ in size and structure from those present in the cytosol. Mitochondria therefore have the genetic information and machinery to synthesise proteins. Mitochondria contain approximately 1000 different proteins, that is, about 3–4% of all cellular proteins. It follows that most mitochondrial proteins are encoded by nuclear genes and are synthesised on ribosomes in the cytosol. These proteins require transportation across one or both mitochondrial membranes to reach their sites of activity.

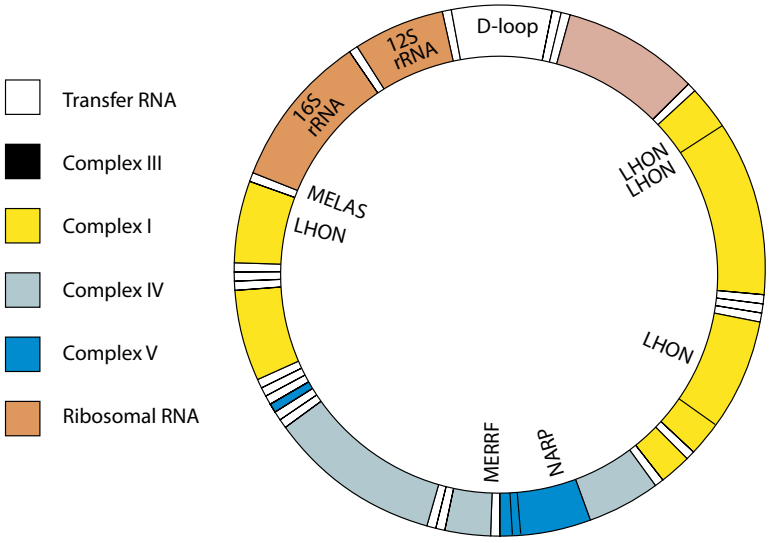


Figure 16.23 Schematic of a human mitochondrial DNA molecule showing its major regions and the sites of mutations associated with some mitochondrial disorders (see *Table 16.6*). This diagram was compiled from several sources. LHON, Leber hereditary optic neuropathy; MELAS, mitochondrial encephalomyopathy, lactic acidosis and stroke-like episodes; MERRF, myoclonic epilepsy and ragged-red fibre; NARP, neurogenic muscle weakness, ataxia and retinitis pigmentosa.

Mitochondria replicate by simple divisions that are independent of mitosis and meiosis (*Chapter 15*). The mitochondrial chromosomes are copied prior to replication. However, the mtDNA-dependent DNA polymerase, or DNA polymerase γ, copies mtDNA with much poorer fidelity than nuclear DNA polymerases. Mutation rates of mitochondrial genes are estimated to be about 10 times greater than those of nuclear genes. This may contribute to the ageing process (*Chapter 18*). Furthermore, the mitochondrial genome resembles that of bacteria in that the genes lack introns and repeat DNA sequences so that about 93% of the DNA is transcribed as opposed to the 5% or less for the nuclear genome. Thus, despite its relatively small size, a relatively large number of mutations are associated with mtDNA and these mutations are significant contributors to human disease.

Causes of mitochondrial disorders

Mitochondrial disorders or myopathies are a heterogeneous group of conditions resulting largely from impairment of the mitochondrial oxidative phosphorylation system. It is estimated they occur with a frequency of up to about 15 in 100 000. Mitochondrial disorders are associated with numerous mutations of the mtDNA but also with mutations in nuclear genes that affect processes such as the assembly of protein subunits and the import of proteins from the cytosol into the mitochondrion. Mutations can affect the morphology of the mitochondrion (*Figure 16.24*). Mitochondrial disorders are

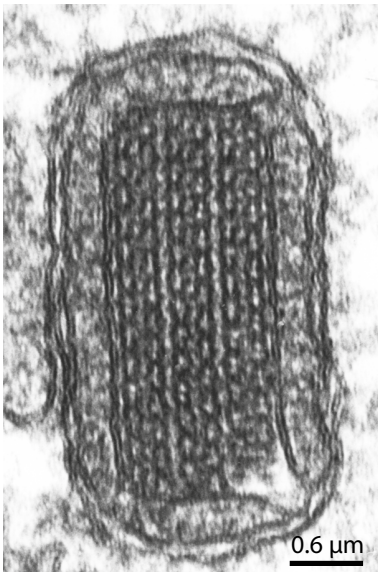


Figure 16.24 Electron micrograph showing the abnormal morphology of a mitochondrion associated with a mitochondrial disorder. Compare with *Figure 16.22*. Courtesy of M.J. Cullen, Muscular Dystrophy Research Laboratories, Newcastle General Hospital, UK.

caused by similar types of mutations to those associated with the nuclear genome, that is, point mutations, deletions and duplications (*Chapter 15*). Interestingly, nuclear mutations are associated mainly with childhood mitochondrial diseases, while adult forms correlate mostly with mtDNA defects. However, the inheritance of the latter differs strikingly from that of nuclear genetic diseases. In zygote formation, a spermatozoon contributes its nuclear genome but not its mitochondrial genome to the egg. The resulting fertilised zygote contains only mitochondria that were present in the unfertilised egg and are therefore maternal in origin. Hence, all children inherit mitochondria only from their mother, and males cannot transmit their mtDNA to subsequent generations (*Figure 16.25*). A typical disease resulting from a mutation in mtDNA is an inherited condition that can affect both sexes but can only be passed on by affected mothers.

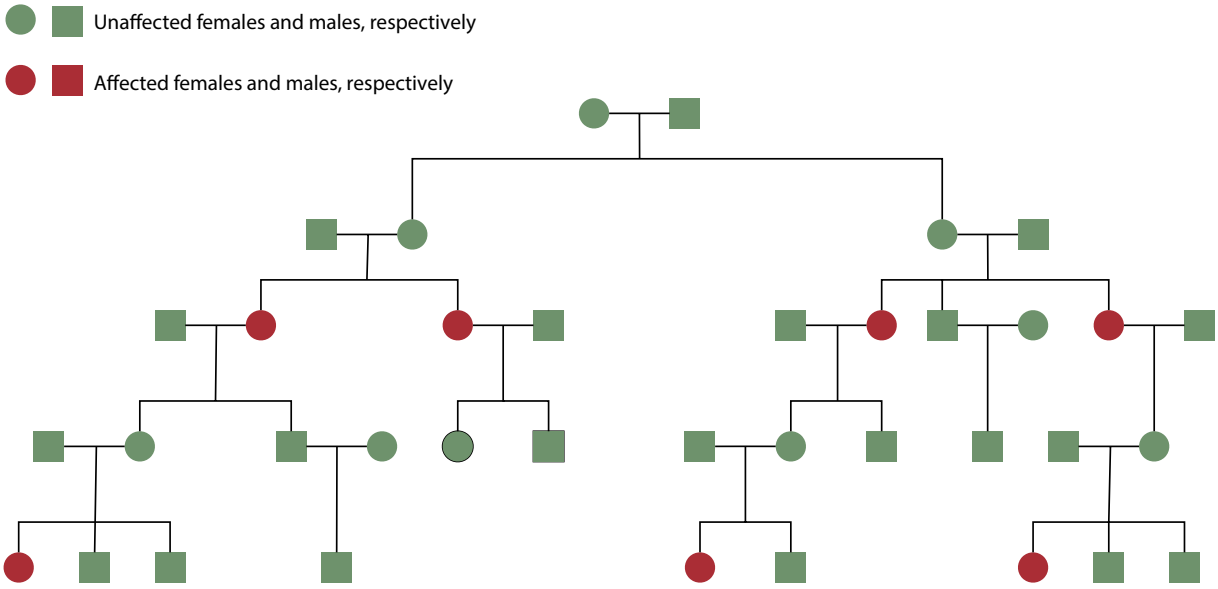


Figure 16.25 An outline of the mode of inheritance associated with many mitochondrial diseases. Redrawn from Wallis C (1999) *The Genetic Basis of Human Disease*. Portland Press.

The first disease to be directly linked to mutations in mtDNA was Leber hereditary optic neuropathy (LHON), an inherited neuropathology. This disease presents in midlife with a rapid bilateral loss of central vision due to atrophy of the optic nerve leading to blindness. Early studies of LHON showed a puzzling maternal pattern of inheritance, with more males than females affected. In 1988, it was demonstrated that LHON was caused by mutations in the mtDNA. Three main mutations are present in at least 90% of families with the disease and all cause substitutions of highly conserved amino acid residues in the affected mitochondrial proteins (*Figure 16.26*).

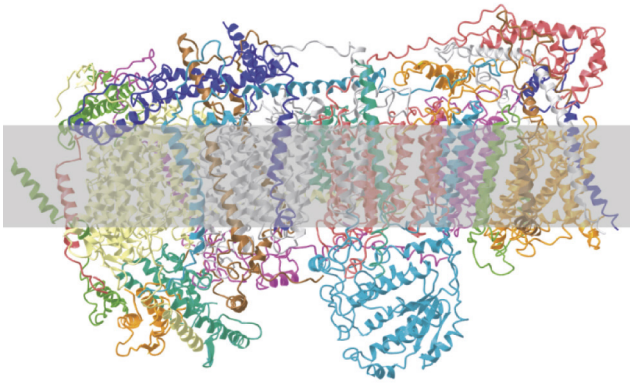


Figure 16.26 Molecular model of human respiratory complex I. Mutations in the genes for subunits of its nicotinamide adenine dinucleotide (NADH), ubiquinone reductase, are responsible for the first identified mitochondrial disease, Leber hereditary optic neuropathy (LHON). The grey block indicates the approximated position of the inner mitochondrial membrane. PDB file 5XTC.

Mutations in mtDNA explain the maternal transmission of LHON, although the reason why a higher proportion of males than females are affected is not known. It has been suggested that the expression of the disease could also require the coinheritance of an X-linked recessive mutation; the fact that the development of LHON could be hormonally influenced by androgens or environmental factors, such as heavy tobacco smoking, may

also contribute to the progression of the condition. Many 100s of pathogenic mutations in mtDNA are associated with a variety of human diseases. However, numerous variants are classified only as 'reported' and await a final confirmation of pathogenicity. However, a number lead to general, multisystem disorders (*Table 16.6*), which, like LHON, are often neuromuscular in nature. Perhaps not surprisingly, most mitochondrial disorders tend to affect the most energy-demanding tissues, such as the central nervous system (CNS), heart and skeletal muscles, kidneys, and secretory tissues, although this is not always the case, since mutant proteins may interfere with mitochondrial functions in addition to ATP formation. The diseases also vary in severity in patients with the same mtDNA mutation. **Homoplasmy** refers to the condition in which every mtDNA molecule has the same causative mutation. In contrast, **heteroplasmy** occurs when the cell has a mixed population of normal and mutant mitochondria. If heteroplasmy occurs, the severity and symptoms of the disease may depend on the proportion of abnormal mtDNA in the critical tissue. The proportion can also differ between mother and baby because of the random segregation of mtDNA molecules at cell division.

Disease	Mutation(s)	General features
Chronic progressive external ophthalmoplegia (CPEO)	Deletion with loss of several genes	Paralysis of eye muscle, myopathy
Kearns–Sayre syndrome (KSS)	Deletion with loss of several genes	CPEO and retinal deterioration, ataxia, heart disease, loss of hearing, kidney failure and diabetes
Leber hereditary optic neuropathy (LHON)	Mutations in genes for subunits of NADH dehydrogenase	Damage to the optic nerve leading to blindness
Leigh syndrome	Mutations in genes for subunits of ATP synthase and inner membrane oxidation complexes	Loss of verbal and motor skills following degeneration of basal ganglia
Mitochondrial encephalomyopathy, lactic acidosis and stroke-like episodes (MELAS)	Mutations in genes for $tRNA^{Leu}$, $tRNA^{Glu}$ and NADH dehydrogenase	Myopathy, dementia, seizures, lactic acidosis
Myoneurogastrointestinal encephalopathy (MNGIE)	Multiple mtDNA deletions	Progressive external ophthalmoplegia, gastrointestinal dysmotility (often pseudo obstruction), diffuse leukoencephalopathy, myopathy
Myoclonic epilepsy and ragged-red fibre (MERRF)	Mutations in genes for $tRNA^{Lys}$ and $tRNA^{Leu}$	Myoclonic seizures, ataxia, lactic acidosis, 'ragged-red fibres' (an abnormality of the tissue when seen under a microscope)
Neurogenic muscle weakness, ataxia and retinitis pigmentosa (NARP)	Mutations in genes for subunits of ATP synthase	Muscle wasting, ataxia, blindness
Pearson syndrome	Deletion with loss of several genes	Childhood bone marrow dysfunction, multiple blood disorders and pancreatic failure

Table 16.6 Examples of mitochondrial diseases. ⏎

BOX 16.2 ACQUIRED MITOCHONDRIAL DNA DEFECTS

The most common cause of acquired mtDNA defects occurs in human immunodeficiency virus (HIV)-1 sufferers treated with highly active antiretroviral therapy (HAART; *Chapter 3*). Highly active antiretroviral therapy for HIV-1 is a combination therapy that uses inhibitors of the viral reverse transcriptase and the HIV protease. The first inhibitor of reverse transcriptase to be used clinically was the nucleoside analogue zidovudine or AZT (see *Figure 3.23*). An unwanted effect of this drug is to inhibit the replication of mtDNA. This depletes the amount of mtDNA and leads to side effects that mimic the symptoms of inherited mitochondrial disorders. Unfortunately, such effects are common to most nucleoside analogue inhibitors of reverse transcriptase. It is now routine practice to monitor mitochondrial functions in patients treated with HAART.

Diagnosis and treatment of mitochondrial disorders

Considerable advances have been made in the understanding of the pathology of mitochondrial disorders. However, an initial diagnosis is problematic because of the varied aetiology and clinical features of these diseases. Similarly, predictors of disease progression are also often unsatisfactory. The concentrations of lactate in plasma, cerebrospinal fluid and urine may, individually or collectively, be increased relative to normal,

although these changes are also seen in numerous other clinical conditions. However, a plasma lactate:pyruvate ratio greater than 40 is usually considered a significant indicator of mitochondrial dysfunction in adults. Other indicators may be the presence of organic acids and myoglobin in the urine, ketoacidosis (*Chapter 7*) and impaired renal, liver and glandular functions. Next-generation DNA sequencing methods (*Chapter 17*) are rapid and cheap; hence, an initial diagnosis is easily confirmed (or rejected) by identifying the specific mtDNA or nuclear mutation.

Most current treatments of mitochondrial disorders are unsatisfactory. The clinical management of patients is largely supportive and concentrates merely on relieving symptoms. For example, administration of hydrogen carbonate and dialysis may be used to relieve lactate acidosis. Strenuous exercise stimulates mitochondrial biogenesis and helps decrease the proportion of mutant mtDNA in muscle. Treatment with vitamins and ubiquinone supplements has been tried but any benefits have yet to be established.

The last 20-or-so years have seen extensive research, largely using mouse models, to try and establish improved treatments. However, the low patient numbers have hampered attempts at randomised controlled drug and genetic therapy trials. Many low-M_r compounds have been or are under investigation. One, idebenone (*Figure 16.27*), is currently used to treat visual impairment in LHON. This drug is an analogue of coenzyme Q and can penetrate the blood–brain barrier. On its reduction in cells, it is able to donate electrons to respiratory complex III, directly bypassing the defective complex I. Molecular biological techniques to, for example, reduce the heteroplasmy, allowing mitochondrial replication to approach homoplasmy, have to date been disappointing. Gene therapy using **allotopic expression** may offer a permanent cure in the future. The strategy behind this technique is to insert a mitochondrial gene with the addition of a mitochondrial targeting sequence into the nucleus, where it is transcribed. The resulting mRNA is translated in the cytosol and the protein product would then be imported into the mitochondria and counter the effects of the mtDNA mutation. Unfortunately, attempts to alleviate LHON using a viral vector to insert the relevant native gene have shown only limited improvement.

Families affected by mitochondrial disorders should be offered counselling on the prognosis, risk of recurrence and prevention. This is complex because causes include nuclear mutations and mitochondrial types that are affected by the degree of mutant mtDNA, which can change dramatically during each transmission from a mother to a child.

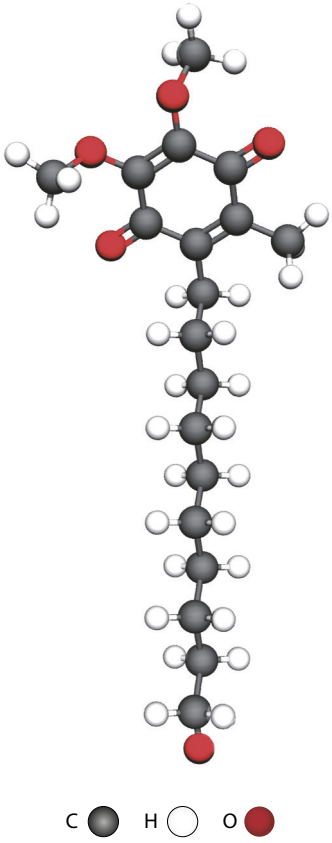


Figure 16.27 Computer drawn model of idebenone.

16.6 PEROXISOMAL DISORDERS

Peroxisomes (*Figure 16.28*), like lysosomes, are limited by a single membrane. They perform certain oxidation reactions, particularly those involved in the partial degradation of very long-chain fatty acids (VLCFAs) and plasmalogens, which are phospholipids but have a fatty aldehyde alcohol rather than a fatty acid attached to their glycerol moiety (*Chapter 10*). Unlike mitochondria, these oxidations are not coupled to the formation of ATP. Matrix (interior) peroxisomal proteins are synthesised in the cytosol and imported into the organelles by a complex process that relies upon the relevant protein possessing a peroxisomal targeting signal (PTS). The major one (PTS1) consists of the carboxyl terminal consensus sequence SKL (-Ser-Lys-Leu-COO^-). Peroxisomal PTS1 is recognised by a receptor encoded by the gene *PEX5*, which imports them into peroxisomes. However, some peroxisomal matrix proteins possess a less conserved PTS2 that is recognised by the receptor PEX7. The term PEX refers to genes identified as necessary for peroxisomal biosynthesis; their numerical attribution, in general, refers to the order in which they were identified.

Defects in individual peroxisomal enzymes or faults in their transport from the cytosol into the peroxisome produce a variety of potentially lethal human diseases. All are rare, occurring in approximately 1 in 50 000 live births. The most common are:

- adrenoleukodystrophy
- Refsum disease
- rhizomelic chondrodysplasia punctata
- Zellweger spectrum disorder

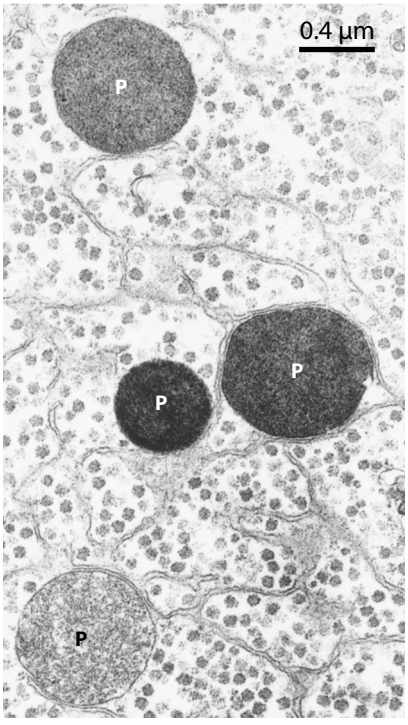


Figure 16.28 Electron micrograph of liver peroxisomes (P) identified by their catalase activity. Note that the intensity of staining varies between individual organelles. Courtesy of Drs M. Espeel and F. Roels, Department of Anatomy, Embryology and Histology, University of Ghent, Belgium.

Adrenoleukodystrophy

Adrenoleukodystrophy (ALD) is an X-linked disease (*Chapter 15*) that affects approximately 1 in 21 000 males. However, a neonatal form of ALD (NALD) that is inherited in an autosomal recessive fashion is now classified as a form of Zellweger spectrum disorder (see below). Adrenoleukodystrophy is associated with mutations in *ABCD1* sited at Xp28, which encodes an ATPase binding cassette similar to CFTR (*Section 16.2*). More

than 700 different *ABCD1* mutations have been noted. Defective gene products result in impaired transport of VLCFAs into peroxisomes where they would normally be subjected to peroxisomal β oxidation. This results in an accumulation of VLCFAs in the plasma and primarily in myelin of the CNS, the adrenal cortex with progressive adrenal cortex dysfunction and Leydig cells of the testes (*Chapter 7*) that can produce one of several types of ALD. The most severe affects males four to eight years of age and causes severe mental impairment. The major symptoms are learning disabilities, seizures, deafness, dysarthria, dysphagia, poor muscular coordination, vomiting, increased pigmentation of the skin, progressive dementia and poor behaviour, such as aggression and hyperactivity. The adult form normally presents at 20–36 years. Symptoms include stiffness and a progressive **paraparesis** or weakness of the legs, and **ataxia**, an inability to control muscular movements. Progress of the adult form is slower than childhood ALD but may also result in a loss of brain function and may be confused with multiple sclerosis. Occasionally, women carriers can show mild symptoms of ALD.

The diagnosis of peroxisomal disorders usually involves biochemical tests, such as determining the concentrations of VLCFAs in the plasma. Observations of deficient secretions of aldosterone and cortisol (*Chapter 7*) by the adrenal cortex can imply ALD. Prenatal diagnoses are available using cultured amniocytes and chorionic villus cells (*Chapter 15*). Patient management largely involves supportive care and symptomatic therapy. The prognosis for childhood ALD is particularly poor, with most sufferers dying before their first birthday. Patients may be given Lorenzo's oil (*Box 16.3*) and a strict dietary regimen that returns VLCFAs to normal concentrations but does not prevent neurological degeneration. Haemopoietic stem cell transplants usually have a successful outcome, but only if diagnosis is early and the donor is a perfect human leukocyte antigen (HLA) match; however, it will not reverse any cerebral damage. Allogenic gene therapy may offer more success in the future (*Chapter 6*).

BOX 16.3 LORENZO'S OIL

Lorenzo's oil gained considerable publicity due to the 1992 film of the same name. The film is based on the story of Augusto and Michaela Odone's efforts to save the life of their son, Lorenzo, who was diagnosed with incurable adrenoleukodystrophy (ALD) in 1984. His parents were, however, unwilling to give up the struggle to save him and, after participating in several failed therapies, began to investigate the disease themselves. This resulted in the treatment of Lorenzo with a mixture of glyceryltrioleate and glyceryltrierucate, hence 'Lorenzo's oil', in the hope that it would reduce the concentration of longer chain fatty acids in the body and also reduce demyelination and clinical progression.

Lorenzo's oil rapidly reduces very long-chain fatty acids in plasma to normal or near-normal levels. Unfortunately, in patients with neurological symptoms, its use does not alleviate symptoms or delay the progression of the disease. Postmortem studies show that very long-chain fatty acids in the brain are apparently unaffected.

Refsum disease

Sigvald Refsum (1907–1991) first recognised the disease (RD) in the 1940s. He observed markedly increased concentrations of phytanic acid (strictly, phytanate; see *Margin Note 8.3*), a major constituent of dairy foods, in certain patients. The disease was previously described as occurring in as two forms: adult (ARD) and infantile (IRD). However, IRD is now incorporated into Zellweger spectrum disorder (see below).

Adult RD is an autosomal recessive disorder, with most cases arising from mutations in *PHYH* sited at 10p13, which encodes α phytanoyl CoA hydroxylase (PAHX; *Figure 16.29*), which decreases the α oxidation of phytanic acid. The peroxisomal enzymes PAHX and α hydroxyphytanoyl-CoA lyase are necessary for the oxidation of phytanic acid (*Figure 16.30*).

Clinical features of RD include an accumulation of phytanic acid in plasma and lipid-containing tissues, retinitis pigmentosa, blindness, **anosmia** (loss of sense of smell), deafness, sensory neuropathy and ataxia. Infantile RD is caused by mutations in genes necessary for peroxisome biogenesis, such as *PEX1* that encodes peroxisome biogenesis factor 1. This also leads to an accumulation of phytanic acid but resembles Zellweger spectrum disorder (see below) in that it is due to general defects in the transport of appropriate peroxisomal enzymes into the organelle. Refsum disease is relatively less severe than Zellweger syndrome because the enzymes in the cytosol are thought to have longer half-lives.

Diagnosing RD is based on the general clinical presentation, determining peroxisomal enzyme activity and evaluating the concentration of phytanic acid in plasma samples. Prognosis for RD varies dramatically. Sufferers are given a dietary regimen

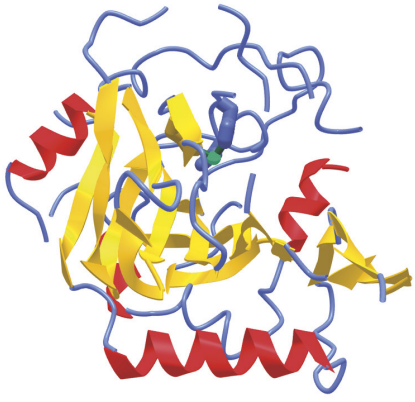


Figure 16.29 Molecular model of α phytanoyl CoA hydroxylase. The green sphere represents its bound Fe^{2+}. PDB file 2A1X.

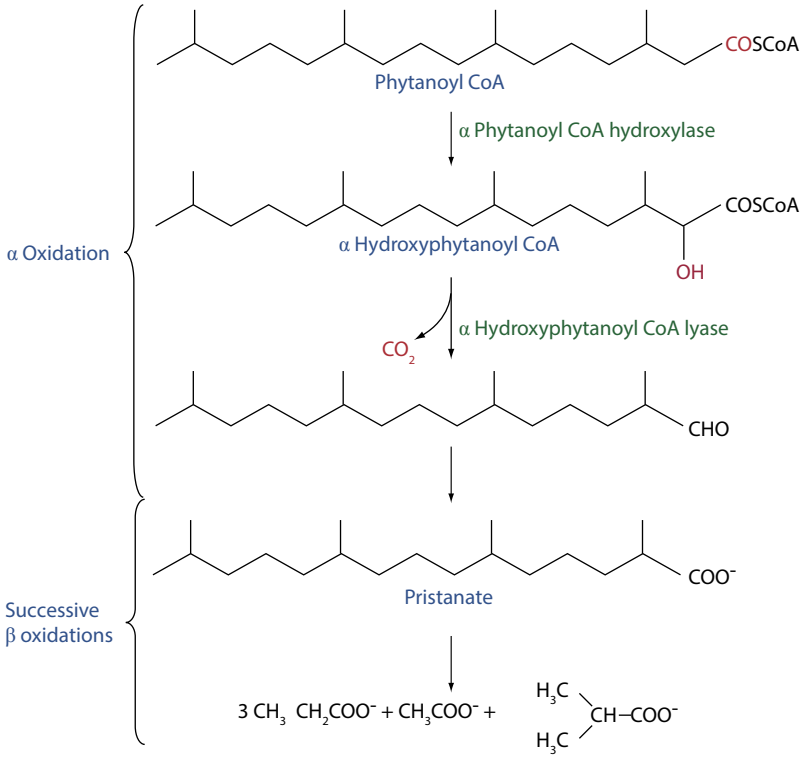


Figure 16.30 Outline of the catabolism of phytanic acid.

free of phytanic acid. Strict dietary adherence can cause the neurological symptoms to arrest, and nerve response can improve. Weakness, problems with walking and scaliness of the skin can also improve. Prior to the initiation of dietary therapy, more than half of patients died before the age of 30 years. However, since the introduction of dietary therapy, very few patients have died so young. Infantile RD may present as young as six months, and life expectancy can extend into the child's teens.

Rhizomelic chondrodysplasia punctata

Rhizomelic chondrodysplasia punctata (RCDP) are disorders of peroxisomal biogenesis. Five types have been noted, each being caused by mutations in different genes (*Table 16.7*). The best characterised is RCDP1, which occurs in severe and mild forms. Severe RCDP1 is characterised by rhizomelia and punctate calcifications of epiphysial cartilage (*Margin Note 16.3*). The birth weights of patients are low, subsequent growth is retarded and cataracts are common. Intellectual development is severely restricted and seizures are common. Unfortunately, most children die before the age of 10 years. Mild RCDP1 presents in early childhood, with the extent of rhizomelia more variable,

Types	Gene (location)	Protein
RCDP1	*PEX7* (6q23.3)	PTS2 receptor (peroxisome biogenesis factor 7 or peroxin 7)
RCDP2	*GNPAT* (1q42.2)	Glyceronephosphate *O* acyltransferase (peroxisome specific enzyme)
RCDP3	*AGPS* (2q31.2)	Alkyldihydroxyacetonephosphate synthase (essential for ether lipid biosynthesis in peroxisomes)
RCDP4	*FAR1* (11p15.3)	Peroxisomal fatty acyl CoA reductase 1 (necessary for biosynthesis of wax esters)
RCDP5	*PEX5* (12p13.31)	PTS1 receptor (peroxisome biogenesis factor 5 or peroxin 5)

PTS, peroxisomal targeting signal.

Table 16.7 Genes and proteins associated with rhizomelic chondrodysplasia punctata (RCDP).

Margin Note 16.3 Rhizomelic chondrodysplasia punctata: What s in a name?

Even by the standards of many medical terms, rhizomelic chondrodysplasia punctata is something of a mouthful. Rhizomelic is derived from the Greek *rhizo* and *melos* meaning root and limb, respectively, and is a generalised term referring to the shoulder and pelvic girdles. Hence, rhizomelia is the severe shortening of long bones. Chondrodysplasia is also derived from Greek, namely *chondros*, *dys* and *plassein* meaning cartilage, bad and to form, respectively, while punctata (Greek *punctum*, point) refers to elevated spots or punctures in a tissue.

cataracts, poor development and joint complications. Most patients can walk and communicate to some extent and survive into childhood.

An initial diagnosis of RCDP1 is based on the pattern of symptoms presented and radiological observations of the skeleton. Confirmatory evidence would include demonstrating decreased concentrations of plasmalogens in erythrocytes but an increased level of phytanate and normal values of VLCFAs in the plasma. Molecular biological techniques, such next-generation DNA sequencing (*Chapter 17*), may be used to identify specific mutations. There is no cure for RCDP1; management is supportive, with treatment aimed at alleviating physical stress, such as physiotherapy, cataract removal to improve vision and an appropriate diet. Genetic counselling for affected families is, of course, desirable.

Zellweger spectrum disorder

Hans Zellweger (1909–1990) spectrum disorder (ZSD), like RCDP, is a series of disorders of peroxisomal biogenesis. Peroxisomal enzymes remain in the cytosol and are rapidly degraded, leading to peroxisomal deficiency and a range of phenotypes, hence the term spectrum. They include Zellweger syndrome (ZS), NALD and IRD, which were individually described before the genetic/molecular bases of the conditions were established. Severe ZSD was previously called ZS. Patients have severe abnormalities in their brain, liver and kidneys and chondrodysplasia punctata (*Margin Note 16.3*). Sadly, most babies die within a year of birth. Intermediate ZSD was previously described as NALD and IRD, although the latter is sometimes assigned to the mild end of the spectrum. The condition can progress with degeneration of eye and ear functions, the myelin of the CNS, the development of ALD and adrenal insufficiency (*Chapter 7*), and death of the patient. In nonprogressive forms, the infant can survive up to school age and show normal intelligence; however, adrenal insufficiency and defective teeth may develop. The diagnosis of ZSD requires care, given the range of conditions, but is based on the clinical presentation, biochemical analyses and molecular biological investigations as described above for other peroxisomal disorders. Treatment is mainly supportive, as in RCDP.

BOX 16.4 GLYCOSOMES AND *PEX*S

Glycosomes are organelles related to peroxisome and peculiar to some trypanosome protozoa, including *Trypanosoma brucei* and *cruzi*, the causative organisms of African trypanosomiasis and Chagas disease, respectively (*Chapters 2 and 3*). They contain glycolytic enzymes, although these differ from the enzymes of glycolysis found in the cytosol of other organisms, and also perform some aspects of intermediary metabolism. The biogenesis of glycosomes depends on interactions between peroxisome biogenesis proteins, notably PEX5 and PEX14. Several molecules of low M_r have been shown to inhibit interactions between these peroxins in *T. brucei* (*Figure 16.31*) resulting in, for example, glucose toxicity, the depletion of ATP and the death of the parasites. It is hoped that this will form the basis of new therapies to treat these widespread illnesses.

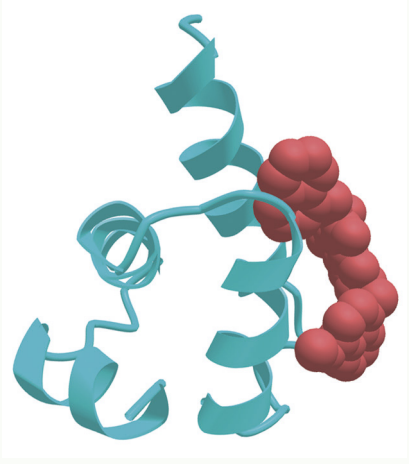


Figure 16.31 Molecular model showing the binding of a molecule of low M_r (red) to PEX14. This binding prevents interactions between PEX5 and PEX14. PDB file 5N8V.

16.7 CYTOSKELETAL DISORDERS

The cytoplasm of nucleated cells is supported by a cytoskeleton consisting of fibrous and their associated proteins. The three types of fibres are:

1. microfilaments
2. intermediate filaments
3. microtubules

All three fibres are actively extended and shortened during cellular activities so their lengths are indeterminate but their diameters differ and are fairly uniform between cell types. Microfilaments (MFs) have a diameter of approximately 7 nm and microtubules (MTs) about 25 nm, while intermediate filaments (IFs) have diameters of 8–11 nm, hence their respective names.

Microfilaments are relatively flexible filaments composed of bundles of cross-linked actin proteins that can withstand compression. Microtubules are formed of tubulin proteins arranged into hollow rods that are rigid and can resist both compression and tension. Intermediate filaments are built of several types of tissue-specific proteins, for example keratins in epidermal cells and desmin in muscles. They form flexible cables whose high tensile strength allows the cell to resist excessive stretching (*Figure 16.32*). Microfilaments and MTs can form defined transport tracks within the cell. The movements of macromolecules and membranous structures along these tracks involves interactions with motor proteins (dyneins and kinesins in the case of MTs and myosins with MFs), which use free energy from the hydrolysis of ATP to drive transport.

Complexes of actin–myosin are probably best known as the contractile apparatus of skeletal muscle tissue. Skeletal muscle tissue shows a multinuclear organisation or syncytium arranged into fibres, which are surrounded by a basal lamina of extracellular matrix proteins that forms a supporting sheath. Each fibre contains sarcoplasm (cytoplasm) that houses the actin–myosin contractile fibres and is surrounded by a sarcolemma (plasma membrane). Dystrophin (*Figure 16.33*) forms a network of elongated molecules about 150 nm long within the sarcoplasm. This protein links actin filaments to a transmembrane complex of proteins that, in turn, is linked to components of the basal lamina (*Figure 16.34*). This complex arrangement of proteins provides mechanical support to the sarcolemma during muscle contraction.

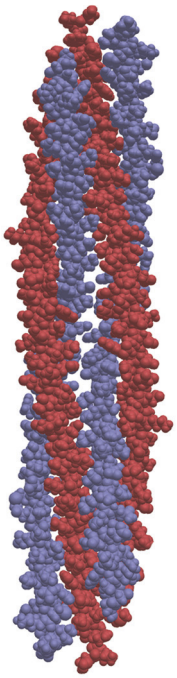


Figure 16.32 Molecular model of a portion of an intermediate filament consisting of two coiled coils (see *Margin Note 16.2*). PDB file 6ECO. ⏎

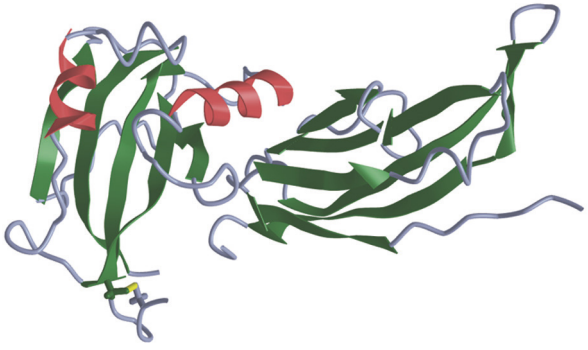


Figure 16.33 Molecular model of a dystrophin subunit. PDB file 5LLK. ⏎

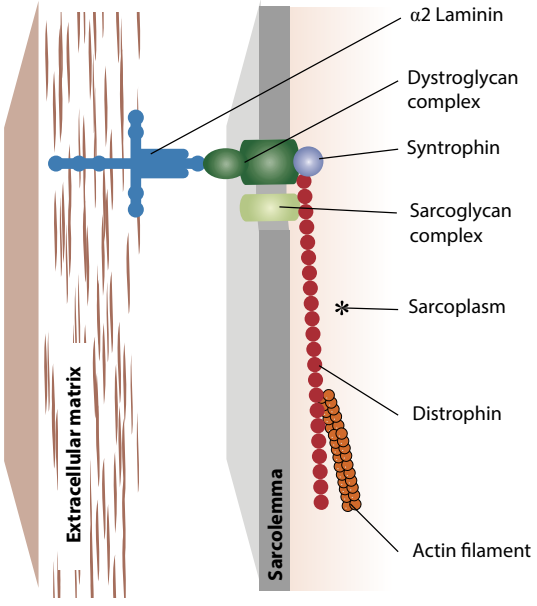


Figure 16.34 Simplified schematic indicating how dystrophin, a transmembrane protein complex, and the basal lamina interact to stabilise the sarcolemma of muscle fibres. ⏎

The unique biconcave shape of erythrocytes (*Chapter 13*) is maintained by a specialised cytoskeleton composed of five major proteins that form a network lining the inner surface of their plasma membranes (*Figure 16.35*). A spectrin–actin complex is thought to function in a manner that resembles that of the dystrophin–actin complex of skeletal muscle, providing mechanical support to the plasma membrane and preventing its rupture during circulation. The flexibility of the network also allows erythrocytes to deform and spring back into shape as they are pumped through narrow capillaries.

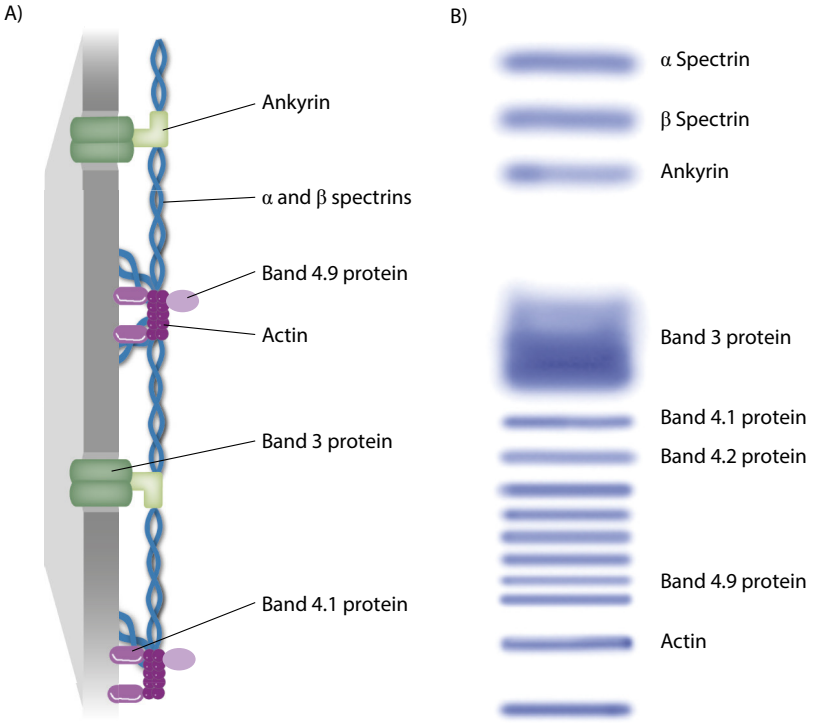


Figure 16.35 (A) Simplified schematic of the cytoskeleton of an erythrocyte. (B) Separation of erythrocyte cytoskeleton proteins by sodium dodecyl sulphate polyacrylamide gel electrophoresis. ⏎

Muscular dystrophies

Dysfunctional forms of dystrophin, its associated transmembrane complex proteins, laminins and proteins involved in their production can all lead to genetic disorders called muscular dystrophies (MDs; *Table 16.8*). The incidences of MDs vary depending on the specific type. Duchenne MD (DMD) has an incidence of 3 in 10 000 live-born males; Becker MD (BMD), the second most common form, is 1 per 30 000. Other types of MD are much rarer; limb-girdle MDs occur in only about 1% of patients with MDs.

Muscular dystrophy	**Gene (site)**	**Protein(s) involved (see *Figure 16.34*)**
Becker muscular dystrophy (BMD)	*BMD* (Xp21.2-p21.1)	Dystrophin
Congenital muscular dystrophy (CMD)	*LARGE* (22q12.3)	*N* acetylglucosaminyltransferase
Duchenne muscular dystrophy (DMD)	*DMD* (Xp21.2-p21.1)	Dystrophin
Limb-girdle muscular dystrophies (LGMDs)	*SGCA* (17q21.33)	α Sarcoglycans[a]
	SGCB (4q12)	β Sarcoglycans
	SGCG (13q.12.12)	γ Sarcoglycans[b]
	SGCD (5q33.2-q33.3)	δ Sarcoglycan

[a] Mutations to several genes for proteins other than sarcoglycans are also known to cause LGMDs.
[b] Also known as severe childhood autosomal recessive muscular dystrophy (SCARMD).

Table 16.8 Examples of muscular dystrophies. ⏎

Mutations involving the genes for dystrophin and its associated proteins compromise the mechanical strength of muscle fibres leading to their rupture and death during years of contraction; hence, MDs are characterised by a progressive wasting of muscle tissues.

The best described MDs are DMD (also called pseudohypertrophic muscular dystrophy) and BMD (also known as benign pseudohypertrophic muscular dystrophy), which were first described by the French neurologist Guillaume-Benjamin-Amand Duchenne (1806–1875) and the German doctor Peter Becker (1908–2000), respectively. They are caused by different mutations in *DMD* (*Table 16.8*), which encodes dystrophin. Mutations that produce STOP codons or deletions of even relatively small portions that change the reading frame lead to DMD. Surprisingly, BMD is the result of much larger deletion mutations, but the reading frame of the rest of the gene is unaltered, leading to the production of defective dystrophin. This, apparently, is sufficient to protect the muscles from degenerating as badly or as quickly as is the case with DMD. Both DMD and BMD are inherited in an X-linked recessive manner, although up to 30% of cases of DMD are the result of a *de novo* mutation. Like all X-linked recessive characters, the presence of the gene in a male will lead to the development of symptoms because of the absence of a homologous, normal gene on the Y chromosome (*Chapter 15*). Heterozygous females will be carriers and generally symptomless, although a number do show mild symptoms. Only in the unlikely event of a female being a homozygote, with two X chromosomes carrying a mutated gene, will females present with the disease.

Wasting of muscle tissues in DMD is fairly predictable in that it follows an aggressive and progressive course. It usually begins in the legs and pelvis and progresses to the shoulders and neck followed by the respiratory muscles. Symptoms usually appear in male children aged one to six years. Boys with the disorder are often late in learning to walk. Toddlers may present with enlarged calf muscles and a clumsy, unsteady gait. The initial enlarged muscle mass of the calf muscles is a compensation for loss of muscle strength. However, the enlarged muscle tissue is eventually replaced by fat and connective tissue (pseudohypertrophy). Muscle contractures occur in the legs, rendering the muscles unusable, as muscle fibres shorten and fibrosis occurs in connective tissue. Intellectual impairment may also occur, although it is not inevitable and does not worsen as the disorder progresses. At school age, there is trouble climbing stairs and the child may walk on his toes or the balls of his feet and can easily fall over. Characteristically, the belly is stuck out and shoulders held back to maintain balance. Nearly all children with DMD lose the ability to walk between 7 and 12 years and, by 10 years, leg braces may be required for walking. Bones develop abnormally, leading to skeletal deformities of the spine and other areas. In the early teens, or earlier, the heart and respiratory muscles may also be affected. Cardiomyopathy is commonly present, but signs of congestive heart failure or arrhythmias (*Chapter 14*) are rare. The muscular weakness and skeletal deformities frequently lead to breathing disorders, including pneumonia. The aspiration of food or fluid into the lungs is common during the later stages. If untreated, death usually occurs by the age of 25 years, typically from these respiratory disorders.

The symptoms of BMD are considerably milder than those of DMD. They usually begin about the age of 12 years or in early adulthood. The pattern of development resembles that of DMD, although the course is slower and far less predictable. Generally, muscle wasting begins in the legs and pelvis and then progresses to the muscles of the shoulders and neck, followed by loss of arm muscles and respiratory muscles. Pseudohypertrophy of the calf muscle is again apparent. The rate of muscle tissue degeneration in BMD varies greatly. Some men require wheelchairs by their 30s, while others manage for many years with minor aids, such as walking sticks, and, indeed, the ability to walk may continue to the age of 40 or over. Cardiomyopathy is less common than with DMD. If it does occur, then, like DMD, congestive heart failure or arrhythmias are rare. Death usually occurs in the fifth decade, but some patients do live to an advanced age.

Diagnosis, management and treatment of muscular dystrophies

Diagnosing MDs begins with a family history and physical examination. The pattern of weakness can be particularly helpful, and usually the source of weakness, whether in motor neurons or muscle tissue, can be identified by physical examination. Nerve conduction studies can indicate whether the underlying problem is muscular or nervous. An electromyography (EMG) to measure electrical activity in muscles can show if the weakness is caused by destruction of muscle tissue. Another early diagnostic test is to measure creatine kinase (CK) activity of the blood to assess any leakage of the enzyme from damaged muscle (*Figure 16.36*). An increase in CK levels is seen in MDs but also in other conditions, such as an inflammation. A magnetic resonance imaging (MRI; see *Chapter 18*) scan can be useful to visualise intramuscular activity. An examination of

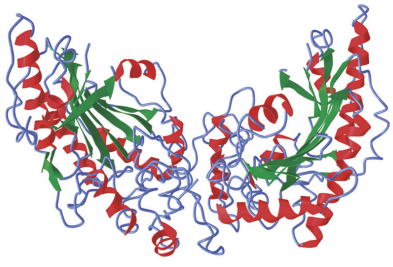


Figure 16.36 Molecular model of skeletal muscle creatine kinase. PDB file 1I0E.

muscle biopsy material can distinguish MDs from inflammatory and other disorders, as well as identify the type of dystrophy present. Genetic testing can confirm a diagnosis using DNA of blood or muscle cells and give precise clinical information. However, a diagnosis of BMD is often not made until adolescence or even adulthood, when physical activities become more difficult.

There are no cures for DMD or BMD; however, optimal care of patients requires a multidisciplinary team. In both diseases, the goal of treatment is to control the symptoms to promote the quality of life. Inactivity, such as prolonged bed rest, generally worsens muscle disease and so physical activity is encouraged. Physiotherapy may be helpful in maintaining muscle strength and coordination. Orthopaedic appliances, for example leg braces, walking sticks and wheelchairs, may improve mobility and the ability for self-care. Glucocorticoids, such as prednisolone (see *Figure 6.29*) and deflazacort therapy, improve muscle activity and delay the development of associated respiratory complications. The last decade or so has seen the development of intravenous gene therapy drugs that are synthetic antisense oligonucleotides (*Margin Note 16.4*). For example, eteplirsen (*Figure 16.37*) binds to specific parts of exon 51 of DMD pre-mRNA, which allows the protein-producing machinery of the cell to 'skip' mutated gene variants in this region. The effect is to restore the translational reading frame of dystrophin mRNA and promote the production of a truncated version of the protein that, nevertheless, contains the essential functional regions. Given, the enormous size of DMD and the large number of mutations that can contribute to the development of MDs, gene therapy using synthetic antisense oligonucleotides is only of benefit to 14–25% of patients, although a number of other antisense oligonucleotides are under development. Gene therapy, with the aim of replacing the defective gene with an active version, has the potential to afford a cure.

Margin Note 16.4 ⏎ Antisense oligonucleotides

Antisense oligonucleotides (ASOs) are short, synthetic, single-stranded deoxyribonucleotides that are complementary to the mRNA target. Generally, they are approximately 20 nucleotides long. The function of ASOs is to regulate target gene expression through several distinct mechanisms, including altering splicing and sterically hindering ribosome activities.

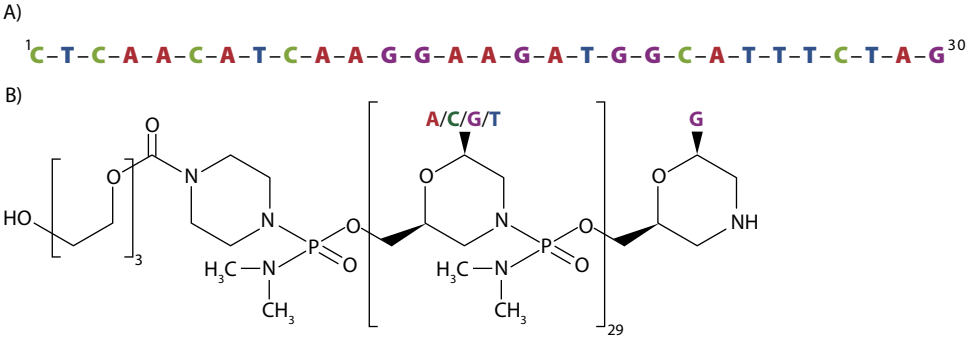


Figure 16.37 (A) The base sequence of and (B) the structure of eteplirsen. ⏎

The stress associated with MDs means that a support group, where members pool common experiences and problems, can be of use. Genetic counselling is advisable, especially if there is a family history of the disorders. Note that sons of a man with BMD (or DMD) will not develop the disease but daughters can be carriers. An amniocentesis with appropriate noninvasive prenatal genetic tests means that DMD is detectable with approximately 95% accuracy during pregnancy.

Hereditary elliptocytosis and hereditary spherocytosis

Two relatively common disorders associated with the erythrocyte cytoskeleton are hereditary elliptocytosis (HE) and hereditary spherocytosis (HS). Their incidences are variable. In the USA, HS is the most common of the hereditary haemolytic anaemias (*Chapter 13*); among people of Northern European descent, its incidence is approximately 1 in 5000, although many regard this as an underestimation. The prevalence of HE is thought to be 1 in 2000–4000. Like sickle cell anaemia (*Chapter 13*), the condition shows a much higher incidence in areas endemic for malaria; in equatorial Africa, for instance, the incidence is approximately 6 per 1000.

Patients with HE or HS biosynthesise lower amounts of erythrocyte cytoskeletal proteins (*Figure 16.35*) than normal or the proteins themselves are defective following mutations in the relevant genes (*Table 16.9*). The misshapen erythrocytes arising from such mutations give the conditions their respective names (*Figure 16.38*), are less deformable than wild types and become trapped in the spleen where they are prematurely degraded. The consequent excessive release of bilirubin leads to jaundice and gallstones (*Chapter 11*). The anaemia caused by reduced numbers of circulating erythrocytes results in the release of immature erythrocytes from the bone marrow.

Condition	Gene (site)	Encoded protein
Hereditary spherocytosis	*ANK1* (8p11.21)	Ankyrin
	EPB42 (15q15.2)	Protein band 4.2 (see *Figure 16.35*)
	SLC4A1 (17q21.31)	Anion exchanger/band 3 (see *Figure 16.35*)
	SPTA1 (1q23.2)	α Spectrin
	SPTB (14q23.3)	β Spectrin
Hereditary elliptocytosis	*SPTA1* (1q23.1)	α Spectrin
	SPTB (14q23.3)	β Spectrin
	EPB41 (1p35.3)	Protein band 4.2 (see *Figure 16.35*)

Table 16.9 Examples of genes and proteins associated with hereditary spherocytosis and elliptocytosis.

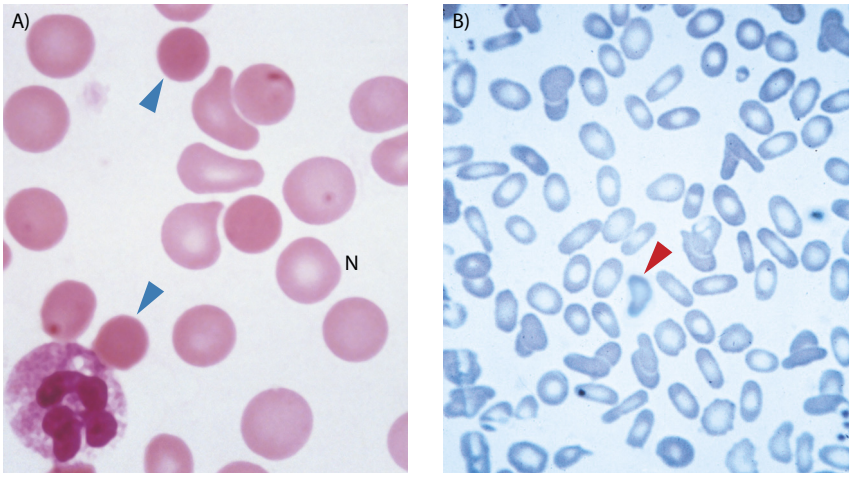


Figure 16.38 Photomicrographs of misshapen erythrocytes in (A) hereditary spherocytosis and (B) hereditary elliptocytosis. In (A), the abnormal spherocytes (examples highlighted) lack the usual biconcave shape (N) and appear a uniform red colour. Note the presence in (B) of abnormal elliptical erythrocytes (example highlighted). Courtesy of A. Will, Royal Manchester Children's Hospital, UK.

Diagnosis and treatment of hereditary elliptocytosis and hereditary spherocytosis

Diagnostic tests for HS and HE include assessing the osmotic fragility of erythrocyte samples to evaluate any increased resistance to osmotic shock, observation of reduced fluorescence following staining with eosin-5-maliemide, and determining any increased amounts of bilirubin and of proteins by electrophoresis in the plasma. The simplest way to test for HE and HS is the microscopic examination of peripheral blood smears to observe the misshapen erythrocytes (*Figure 16.38*).

General treatments include daily ingestion of folate when haemolysis is significant and blood transfusions. Erythropoietin therapy alleviates haemoglobin deficiency and decreases the need for transfusions but increases the requirement for folate. However, the symptoms of HS are alleviated by surgical removal of the spleen (*Table 16.10*). This

	Before splenectomy	After splenectomy
Erythrocyte count $\times 10^{-12}$/L	3.1	5.2
Reticulocytes (%)	16.0	0.95
[Haemoglobin] (g/L)	8.5	15.1
Serum [bilirubin] (mmol/L)	83.0	8.2

Table 16.10 Typical haematological data before and after splenectomy in a patient suffering from hereditary spherocytosis (*see Chapter 13*).

increases the time that the defective erythrocytes stay in the circulation to about 80% of that of a normal erythrocyte. Neonatal sufferers presenting with high concentrations of bilirubin should initially be treated with phototherapy (*Chapter 11*). Most HE patients are asymptomatic, and splenectomy is only required in cases with severe haemolysis. More recently (2024/5), gene therapy is showing some success in treating the condition.

16.8 CASE STUDIES

CASE STUDY 16.1

Bethany, a female infant, suffers from recurrent respiratory infections and presented with clinical evidence of dietary malabsorption. Her brother and sister are asymptomatic. She was sent to hospital for a variety of clinical investigations. After a number of negative tests, she was subjected to a sweat chloride test. The result is given below:

Sweat [Cl^-]	76.0 (normally less than 30 mmol/L)

Questions

(a) Suggest a plausible diagnosis.
(b) Explain how the diagnosis for Bethany was reached.
(c) How should Bethany's siblings be investigated?

CASE STUDY 16.2

Janice is an 18 year old. She was admitted to hospital for clinical tests following a series of frequent heavy nose bleeds and extensive bruising following comparatively minor accidents. She also complained of stiff aching joints and some pains in her long bones. Her erythrocyte and platelet counts were found to be depressed. There was some extension of her abdomen. Routine clinical tests eliminated leukaemia. Enzyme assays showed β glucocerebrosidase activity in leukocytes to be greatly reduced.

Questions

(a) What is Janice suffering from?
(b) What would be the likely progression of this disease if untreated?
(c) What treatment would you recommend for Janice?

CASE STUDY 16.3

Ruth, an eight-month-old baby with a previous medical history of respiratory problems, was brought to her local Accident and Emergency Department with breathing difficulties. Her parents said Ruth had frequent bouts of vomiting, reduced appetite, crying, and poor control of head movements and sucking ability. Ruth's weight was below that expected for her age and gender. She showed **hypotonia** (poor muscle tone), accompanied by muscle weakness and associated ataxia (problems in coordinating movements). Ruth was also failing to achieve accepted developmental targets. Her vision appeared affected with an inability to focus. It was apparent that Ruth had suffered seizures. Ruth's blood oxygen saturation was only 86%. Laboratory tests give a blood lactate concentration of 8.2 (reference value <3 mmol/L) and high levels of pyruvate and alanine, indicating problems in mitochondrial function(s). Magnetic resonance imaging (MRI) and computed tomography (CT) scans of the brain showed neurological lesions. An initial diagnosis was of a neurodegenerative disorder, with the general presentation strongly indicative of Leigh syndrome (LS). Leigh syndrome, Leigh disease or subacute necrotising encephalomyelopathy is an incurable inherited neurometabolic disorder. It is caused by numerous mutations in mitochondrial or nuclear genes. For example:

- *MT-ATP6*, a mitochondrial gene that encodes a subunit of complex V (ATP synthase) of the mitochondrion
- *PDHA1*, located at XP22.12, which encodes a subunit of the mitochondrial matrix, pyruvate dehydrogenase complex
- *SURF1*, which is found at 9q34.2 and encodes an assembly factor for mitochondrial complex IV of the mitochondrial electron transport chain

These mutations result in impaired oxidative metabolism and subsequent neurological deterioration. There was no family history of LS. However, the diagnosis was quickly confirmed by nucleic acid sequencing, which showed mutations in *MT-ATP6*, the most common cause of LS.

Questions

(a) To what class of disorders does LS belong?
(b) Early literature suggested LS was inherited in an autosomal recessive fashion. Explain why this description is inadequate.
(c) Suggest how Ruth might have unfortunately become afflicted with LS.
(d) Mutations associated with LS produce a wide range of clinical manifestations. Suggest why these mutations produce such a drastic presentation.
(e) What steps should be taken in respect of Ruth's family?
(f) Suggest in general terms how LS should be clinically managed.

CASE STUDY 16.4

Colin, a six year old, has a history of adrenal insufficiency. Following a consultation with his general practitioner, he was examined at his local paediatric department following his presentation with muscular weaknesses, impaired balance and partial paralysis of his legs. He was immediately referred to a specialised children's hospital. Clinical tests showed high concentrations of VLCFAs in samples of plasma. A brain scan revealed cerebral atrophy.

Questions

(a) What is the most plausible diagnosis of Colin's condition?
(b) How could this diagnosis be unambiguously confirmed?
(c) Assuming your initial diagnosis is confirmed, what treatment would you recommend?
(d) What other procedures could/should be initiated?

16.9 SUMMARY

Defects in the plasma membrane or those that surround organelles and in the cytoskeleton cause a variety of disorders. Disorders of nucleocytoplasmic transport have been implicated in a number of clinical conditions, including some cancers. Cystic fibrosis is characterised by alterations in the permeability of epithelial cell plasma membranes to Cl^- caused by mutations in *CFTR*. Clinical disorders are also associated with mutated genes that encode nucleoporins and proteins of endomembrane system. Lysosomal storage disorders, of which more than 40 are known, are rare and often fatal, although the adult form of GD is usually associated with a longer lifespan. Mitochondrial diseases are associated largely with defects in the oxidative phosphorylation system, leading to a variety of neurological and neuromuscular disorders. Disorders of peroxisomes result in a number of potentially fatal diseases that are characterised by the accumulation of VLCFAs in tissues. Diseases associated with defects of the cytoskeleton of muscle fibres include DMD and BMD. Mutation in genes for erythrocyte cytoskeleton proteins can lead to them being deformed and to their premature removal by the spleen, with accompanying anaemia.

QUESTIONS

1. Which of the following modes of inheritance is shown by CF?
   (a) Autosomal dominant.
   (b) Autosomal recessive.
   (c) Sex-linked.
   (d) Spontaneous mutation.
   (e) None of the above.

2. Which of the following mutated genes is associated with CF?
   (a) *EPB41*.
   (b) *ANK1*.
   (c) *CFTR*.
   (d) *SPTB*.
   (e) Mitochondrial genes.

3. In relation to organelle-related disorders, which of the following is the odd ONE out?
   (a) LHON.
   (b) AKD.
   (c) RD.
   (d) ZSD.
   (e) RCDP.

4. In treating certain LSDs, explain why it is necessary to use enzyme replacement therapy intravenously but substrate reduction therapy can be given orally.

5. Arrange the following two lists to show the most appropriate pairings:

α Spectrin	LGMD
EPB41	HE
Hexosaminidase A	Sphingomyelin
Sarcoglycans	Cl^- channel
Niemann–Pick disease	MELAS
Refsum disease	BMD, DMD
Adrenoleukodystrophy	Tay–Sachs disease
CFTR	HS
Dystrophin	VLCFAs
Mitochondrial gene for $tRNA^{Leu}$	Phytanic acid

6. Suggest why some mitochondrial diseases lead to myoglobinuria (myoglobin in the urine).

7. Arrange the following two lists to show the most appropriate pairings:

Respiratory complex II	Microtubules
Mitochondrial DNA	Maternal inheritance
Adrenoleukodystrophy	PEX5 and PEX7
Adult Refsum disease	ABCD1 sited at Xp28
Zellweger spectrum disorder	Neonatal adrenoleukodystrophy
Peroxisome biogenesis	Eteplirsen
Dyneins and kinesins	EPB41 at 1p35.3
SCARMD	Succinate: ubiquinone oxidoreductase
DMD	γ Sarcoglycan
Elliptocytosis	α Phytanoyl CoA hydroxylase

8. Anne, a newly born baby, was jaundiced, and after two days showed markedly increased concentrations of bilirubin in her plasma, while a microscopic examination of a blood specimen showed elevated numbers of nucleated erythrocyte precursors and some spherocytes. How should Anne be initially treated?

9. *Figure 16.39* shows the extent of haemolysis of two blood specimens, A and B, in response to decreasing concentrations of NaCl. One specimen was from a patient with hereditary spherocytosis and the other a control.
   (a) Which of the blood erythrocytes shows the greater osmotic fragility?
   (b) Which of the two specimens was from the patient?

10. Gallstones are hard lumps of cholesterol, bile pigments and calcium salts. Explain why patients with HS and HE show an increase in their production of gallstones. You may find it useful to refer to *Chapter 11*.

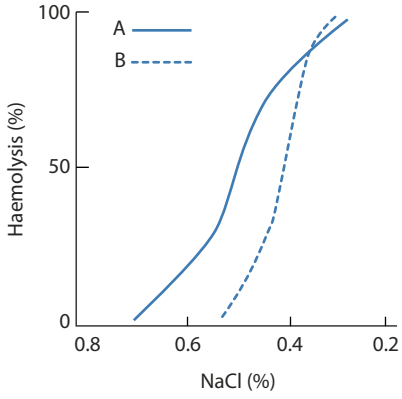


Figure 16.39

FURTHER READING

Braverman NE *et al.* (2016) Peroxisome biogenesis disorders in the Zellweger spectrum: an overview of current diagnosis, clinical manifestations, and treatment guidelines. *Mol Genet Metab* **117**:313–321.

Bris C *et al.* (2018) Bioinformatics tools and databases to assess the pathogenicity of mitochondrial DNA variants in the field of next generation sequencing. *Front Genet* **9**:632. doi: 10.3389/fgene.2018.00632.

Chen Q *et al.* (2021) A review of cystic fibrosis: basic and clinical aspects. *Anim Model Exp Med* **4**:220–232.

Danhelovska T *et al.* (2020) Multisystem mitochondrial diseases due to mutations in mtDNA-encoded subunits of complex I. *BMC Pediatr* **20**:41. doi: 10.1186/s12887-020-1912-x.

Duan D *et al.* (2021) Duchenne muscular dystrophy. *Nat Rev Dis Primers* **7**:13. doi: 10.1038/s41572-021-00248-3.

Enns GM *et al.* (2021) Diagnostic challenges and disease management in patients with a mild Zellweger spectrum disorder phenotype. *Mol Genet Metab* **134**:217–222.

Faria R *et al.* (2023) Delivery systems for mitochondrial gene therapy: a review. *Pharmaceutics* **15**:572. doi: 10.3390/pharmaceutics15020572.

Goetz D and Ren CL (2019) Review of cystic fibrosis. *Pediatr Ann* **48**:e154–e161.

Hachiya N *et al.* (2021) Nuclear envelope and nuclear pore complexes in neurodegenerative diseases – new perspectives for therapeutic interventions. *Mol Neurobiol* **58**:983–995.

Joshua AM *et al.* (2022). Chapter 12: Muscular dystrophies. In: Joshua AM (ed.) *Physiotherapy for Adult Neurological Conditions*. Springer Nature Singapore: Singapore. pp. 731–769.

King JA *et al.* (2022) An update on CFTR modulators as new therapies for cystic fibrosis. *Paediatr Drugs* **24**:321–333.

Lawless C *et al.* (2020) The rise and rise of mitochondrial DNA mutations. *Open Biol* **10**:200061. doi: 10.1098/rsob.200061.

Lin DH and Hoelz A (2019) The structure of the nuclear pore complex (an update). *Annu Rev Biochem* **88**:725–783.

Liu J *et al.* (2021) The role of the Golgi apparatus in disease. *Int J Mol Med* **47**:38.

Liufu T and Wang Z (2020) Treatment for mitochondrial diseases. *Rev Neurosci* **32**:35–47.

Lopes-Pacheco M (2019) CFTR modulators: the changing face of cystic fibrosis in the era of precision medicine. *Front Pharmacol* **10**:1662. doi: 10.3389/fphar.2019.01662.

Manciu S *et al.* (2017) Hereditary spherocytosis – diagnosis, surgical treatment and outcomes. A literature review. *Chirurgia (Bucur)* **112**:110–116.

Marques AR *et al.* (2019) Lysosomal storage disorders – challenges, concepts and avenues for therapy: beyond rare diseases. *J Cell Sci* **132**(2):jcs221739.

Mercuri E *et al.* (2019) Muscular dystrophies. *Lancet* **394**:2025–2038.

Miller W (2017) Stem cell-transplantation therapy for adrenoleukodystrophy: current perspectives. *J Neurorestoratology* **5**:5–19.

Muschalik N and Munro S (2018) Golgins. *Curr Biol* **28**:R374–R376.

Oakes SA and Papa FR (2015) The role of endoplasmic reticulum stress in human pathology. *Annu Rev Pathol* **10**:173–194.

Panteliadis CP and Hagel C (2022) Chapter 10: Refsum disease (heredopathia atactica polyneuritiformis). In: Panteliadis CP *et al.* (eds), *Eurocutaneous Disorders*, 3rd edition. Springer Cham. pp. 445–450.

Perkins HT and Allan V (2021) Intertwined and finely balanced: endoplasmic reticulum morphology, dynamics, function, and diseases. *Cells* **10**:2341–2381.

Rasika S *et al*. (2018) Golgipathies in neurodevelopment: a new view of old defects. *Dev Neurosci* **40**:396–416.
Redhai S and Boutros M (2021) The role of organelles in intestinal function, physiology and disease. *Trends Cell Biol* **31**:485–499.
Ren J *et al*. (2021) Endoplasmic reticulum stress and unfolded protein response in cardiovascular diseases. *Nat Rev Cardiol* **18**:499–521.
Russell OM, *et al*. (2020) Mitochondrial diseases: hope for the future. *Cell* **181**:168–188.
Ryzhkova AI *et al*. (2018) Mitochondrial diseases caused by mtDNA mutations: a mini-review. *Ther Clin Risk Manag* **14**:1933–1942.
Saudou F and Humbert S (2016) The biology of huntingtin. *Neuron* **89**:910–926.
Schuller AP *et al*. (2021) The cellular environment shapes the nuclear pore complex architecture. *Nature* **598**:667–671.
Shteinberg M *et al*. (2021) Cystic fibrosis. *Lancet* **397**:2195–2211.
Spead O *et al*. (2022) Nuclear pore dysfunction in neurodegeneration. *Neurotherapeutics* **19**:1050–1060.
Sprenkle NT *et al*. (2017) Endoplasmic reticulum stress and inflammation in the central nervous system. *Mol Neurodegener* **12**:42–60.
Stirnemann J *et al*. (2017) A review of Gaucher disease pathophysiology, clinical presentation and treatments. *Int J Mol Sci* **18**:441–471.
Tinker RJ *et al*. (2021) Current and emerging clinical treatment in mitochondrial disease. *Mol Diagn Ther* **25**:181–206.
Topaloglu H (2021) Duchenne muscular dystrophy: a short review and treatment update. *Iran J Child Neurol* **15**:9–15.
Turk BR *et al*. (2020) X-linked adrenoleukodystrophy: pathology, pathophysiology, diagnostic testing, newborn screening and therapies. *Int J Dev Neurosci* **80**:52–72.
Valiente-Pallejà A *et al*. (2022) Comprehensive summary of mitochondrial DNA alterations in the postmortem human brain: a systematic review. *EBiomedicine* **76**:103815. doi: 10.1016/j.ebiom.2022.103815.

Useful websites

https://bestpractice.bmj.com/topics/en-gb/1143
https://adc.bmj.com/content/89/9/809
http://www.ncbi.nlm.nih.gov/omim
https://www.hgvs.org/mitochondrial-mutation-databases
https://www.mitomap.org/foswiki/bin/view/MITOMAP/WebHome

Taylor & Francis
Taylor & Francis Group
http://taylorandfrancis.com

Cancer

17

OBJECTIVES

After studying this chapter you should be able to:

- explain what is meant by the following terms: cancer, benign tumour, malignant tumour, metastasis, neoplasm and oncology
- outline the major hallmarks of cancer
- discuss the incidence and epidemiology of cancer
- describe and explain some of the environmental causes of cancer
- outline the relationship between gene mutations and cancer
- describe how cancer may be clinically detected
- describe the clinical features of some cancers
- explain how cancers may be treated

17.1 INTRODUCTION

In multicellular organisms, cells grow by division when stimulated with the requisite internal and external signals (*Figure 17.1*). This growth is rigorously controlled (*Chapter 15*) to ensure that division occurs at an appropriate rate. In some tissues, for example stem cells in bone marrow, the germinative layer of the skin and the epithelium of the gastrointestinal tract (GIT), cells divide constantly and replace ones that have died. In contrast, in nervous tissue, mitosis is rare and lost cells are not replaced. In yet other tissues, cells can commence dividing when required, for example in regenerating liver or during the healing of wounds in the skin.

Cancers are a spectrum of diseases. Their cells (*Figure 17.2*) fail to respond to the usual controlling signals and their growth becomes uncontrolled. Indeed, the name cancer comes from a Latin word meaning 'a crab' and describes the manner in which the pattern of penetration into normal tissues by the abnormal growth bears a superficial resemblance to a crab claw. These abnormal cells may invade nearby tissues and may enter the blood and lymph systems and spread to remoter areas.

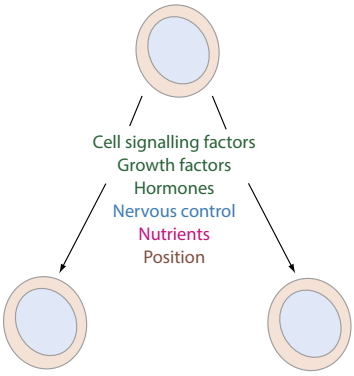


Figure 17.1 Some intrinsic and extrinsic factors that influence cell division.

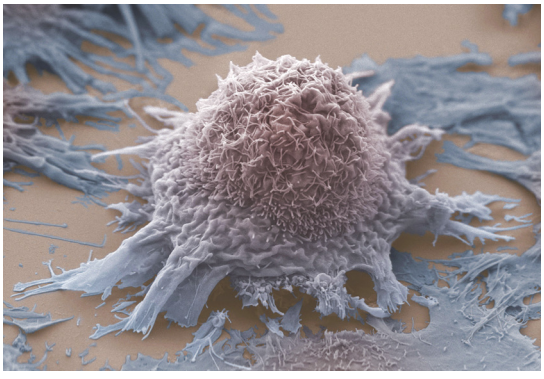


Figure 17.2 False colour electron micrograph of a lung cancer cell. Courtesy of the Wellcome Collection/Anne Weston.

DOI: 10.1201/9781003296836-17

The term cancer is used popularly to describe a **malignant tumour**. A **tumour** is an abnormal mass of cells that may be benign or malignant. *Benign* tumours generally grow slowly and do not spread to other tissues. However, they may continue to grow *in situ* and cause severe clinical problems if they interfere with the normal function(s) of the tissue or if they apply pressure by growing within a confined space, such as in the cranium. A malignant tumour is one that spreads through the blood and lymph circulations from its initial site, where it is known as the **primary tumour**, to establish **secondary tumours** in other tissues and organs (see *Figure 1.6*). The process by which cancer cells migrate from the primary tumour and spread to form secondary tumours is **metastasis**.

A term often used in the context of cancer biology is **neoplasm**. This means, literally, a new tumour or new mass, but is generally used to describe a cancerous growth. **Oncology** (from the Greek *oncos* meaning a lump) is the branch of medicine involved in the study of the development of tumours and their epidemiology, diagnosis and treatment.

In the early 2000s, Robert Weinberg (1942–) and Douglas Hanahan (1951–) proposed that, when cells progress towards a neoplastic state, they acquire distinctive capabilities. These were termed the hallmarks of cancer. The identified hallmarks have expanded in the last 20 years and form a useful framework to help understand tumour pathogenesis (*Table 17.1*).

The causes of cancer are complex and varied. Most cancers arise, ultimately, from mutations in DNA. These mutations may be caused by environmental agents called **carcinogens**, brought about by oncogenic viruses or inherited in the germ line, making individuals more susceptible to cancer.

Cancers can arise from any tissue in the body; indeed, they have been detected in over 200 different sites. However, some sites are more susceptible than others, the most common being the breasts, GIT, lungs and prostate gland. In 2024, the World Health Organization (WHO) noted that cancers are one of the major causes of deaths worldwide, killing 9.7 million people in 2022 or nearly 1 in 9 men and 1 in 12 women.

While most cancers occur more frequently in older rather than in younger people, with cancer generally being regarded as a disease of ageing (*Chapter 18*), certain cancers occur typically in children. In the UK, it has been estimated that one in two individuals will develop a cancer at some time in their life, and that cancers cause one in four deaths.

This chapter will review the biology of cancer and its consequences for the individual. The causes of cancer, including genetic aspects, environmental insults and viruses, will be introduced. The involvement of the pathology laboratory in the screening, diagnosis and treatment of cancers will also be discussed.

Sustained proliferative signalling
Evading growth suppressors
Resisting cell death (apoptosis; see *Box 17.3*)
Replicative immortality
Genome instability and mutations
Inducing angiogenesis
Active invasion and metastasis
Deregulating cellular metabolism
Initiating tumour-promoting inflammation
Avoiding immune destruction

Table 17.1 The 10 major hallmarks of cancers.

Margin Note 17.1
DNA

Germ line DNA refers to the DNA present in the cells that give rise to the gametes, that is, spermatozoa and ova (*Chapter 7*). The sperm and ovum fuse to form a zygote (*Chapters 7* and *15*) and, as mitotic divisions occur, the DNA is inherited by all the cells in the developing embryo. Mutations that occur in germ line DNA are present in the gametes and in all the cells of the individuals to which they give rise.

17.2 CLASSIFICATION OF CANCER TYPES

Cancers are often classified according to their tissue of origin:

- carcinomas
- sarcomas
- leukaemias
- lymphoma
- mixed types and blastomas

Carcinomas, such as adenocarcinomas and squamous cell carcinomas, are derived from epithelial tissue (*Figure 17.3*). Carcinomas account for 80–90% of all cancer cases. A sarcoma is derived from tissues of mesodermal origin, for example muscle, bone, cartilage or fatty tissues. Examples of sarcomas include rhabdomyosarcomas (skeletal muscle), leiomyosarcoma (smooth muscle), osteosarcomas (bone), chondrosarcomas (cartilage), liposarcomas (adipose tissue), gliomas or astrocytomas (neurogenic connective tissue found in the brain) and mesotheliomas (membranous linings of body cavities).

Leukaemia (from the Greek meaning white blood) refers to tumours of the bone marrow, which result in an excess of a single type of cell appearing in the blood. The disease is often associated with the overproduction of immature leukocytes (*Chapters 4* and *13*). Immature leukocytes function poorly; therefore, patients are often prone to recurrent infections. Leukaemia also affects erythrocytes, for instance polycythaemia vera or erythema. Lymphoma refers to a tumour that arises from lymphoid tissue, such as a lymph node. Lymphomas are subclassified into two categories: Hodgkin lymphoma and non-Hodgkin lymphoma. Some tumours may show more than one histological type; these are categorised as mixed type, such as carcinosarcoma and teratocarcinoma.

The highly malignant tumours that occur in childhood all have the suffix *blastoma*. A blastoma is a type of cancer caused by malignancies in precursor cells, which are commonly referred to as blasts. These include neuroblastomas, which occur in the neuroblasts of the adrenal medulla; retinoblastomas, which originate in the retina of the eye; and nephroblastomas, derived from the embryonic cells of the kidney.

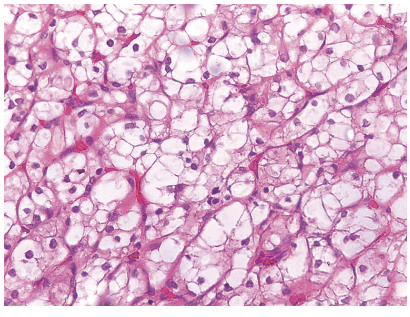


Figure 17.3 Photomicrograph showing a clear cell renal cell carcinoma derived from epithelial tissue.

17.3 EPIDEMIOLOGY OF CANCER

In 2022, 346 217 new cases of malignant cancer were diagnosed in England, with the number of cases in men being slightly higher than that in women. However, individual cancers are often unevenly distributed between the sexes and, clearly, some cancers occur only in men or in women. Obvious examples are testicular and prostate cancers in men and cancer of the *cervix uteri* or 'cervical cancer' in women. Contrary to common belief, men can develop breast cancer, although the incidence is much lower than in women; only approximately 0.5–1% of all breast cancers occur in men. Breast cancer caused 670 000 deaths globally in 2022. The most common forms of cancer in men, women and children globally are shown in *Table 17.2*.

Men	Women	Children
Colorectal cancer	Breast cancer	Leukaemias
Lung cancer	Lung cancer	Lymphomas
Prostate cancer	Colorectal cancer	Brain and other central nervous system tumours

Table 17.2 The most common forms of cancer affecting the world population.

The global numbers of new cases and deaths in 2020 due to different types of cancers is shown in *Figure 17.4*. Except for childhood cancers, the incidence of cancer increases with age. In the UK, the highest incidences occur in 85- to 89-year-old women and men. However, the incidences are significantly higher in men than women in the older age groups (see *Figure 18.21*). This increasing incidence is due to several factors, including increased length of exposure to environmental agents associated with cancer and an accumulation of mutations (*Chapter 18*), coupled with decreased efficiency of the cellular DNA repair mechanisms. In addition, the immune system, which may play a role in eliminating early cancerous cells, also decreases in efficiency with age and this may lead to a failure to eliminate malignant cells as they arise.

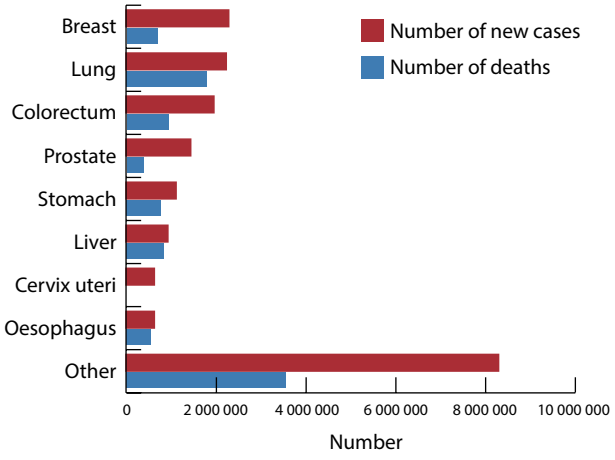


Figure 17.4 The global number of new cases and deaths due to different types of cancers in both female and male patients of all ages. Statistics courtesy of GLOBOCAN (2020).

17.4 MOLECULAR BIOLOGY OF CANCER

Cancers are genetic diseases, that is, they are caused by changes in the DNA of the cell. Mutations, often called variants in clinical circles, such as point mutations, insertions, deletions, inversions, amplifications and chromosomal rearrangements (*Chapter 15*), are common causes of cancer. However, some cancers are not associated with mutations but with epigenetic changes. There are several types of epigenetic changes, including DNA methylation, histone modification and RNA interference. Unlike gene mutations/variants, epigenetic changes do not affect genomic DNA sequences, but can change gene activities.

Margin Note 17.2
Somatic mutation theory

In 1914, Theodor Boveri (1862–1915) suggested that malignant tumours arose from a cell that had acquired chromosomal abnormalities, although the role of chromosomes in inheritance was not known at that time. He also predicted chromosomal (genomic) instability to be a key hallmark of cancers.

The multistep theory of cancer suggests that a succession of genetic alterations must occur before a cancer can develop. The steps of tumorigenesis (*Figure 17.5*) include:

- initiation
- promotion
- progression

Initiation is the induction of a mutation in a critical gene involved in the control of cell proliferation or DNA repair. Mutagenic agents that predispose cells to develop tumours are called *initiators*. The effects of initiators are irreversible. Once mutated by an initiator, the cell is susceptible to the effects of *promoters*. Promoters induce second mutations or epigenetic changes that confer on the cell the ability to grow at a faster rate than would be normal. The cell can then proliferate to form a tissue, which is described as **hyperplastic**. Hyperplastic daughter cells all contain the mutation formed by the initiator. Daughters cell can then undergo further mutations to form a mass of cells that have an abnormal morphology, such as a carcinoma *in situ* of the skin. Accumulation of further genetic and epigenetic alterations within the clone may mean that the cells can detach from the tumour and invade surrounding tissues, that is, they metastasise.

In 2018, the Catalogue of Somatic Mutations in Cancer (COSMIC) Cancer Gene Census (CGC) listed 719 genes, forming more than 3% of the total number of human genes (ca 22 000), that are associated with cancers. Approximately 90% of these somatic mutations were detected in cancer cells. Approximately 20% of 'cancer genes' show mutations in the germ line DNA that predispose an individual to cancer, while 10% are associated with mutations in both germ line and somatic DNA. In 2020, a Pan-Cancer Analysis of Whole Genomes (PCAWG) of 2583 cancer patients showed that 91% of tumours had at least one driver mutation. There is considerable heterogeneity, with an average of four to five driver mutations per tumour somatic cell in both cancer patients and the types of tumours, including single-nucleotide variations (SNVs), multinucleotide variants, insertions and deletions (indels), copy-number alterations, retrotransposition and mitochondrial DNA mutations. Approximately 9% bear point mutation in the telomerase reverse transcriptase (*TERT*) promoter (*Box 17.1*).

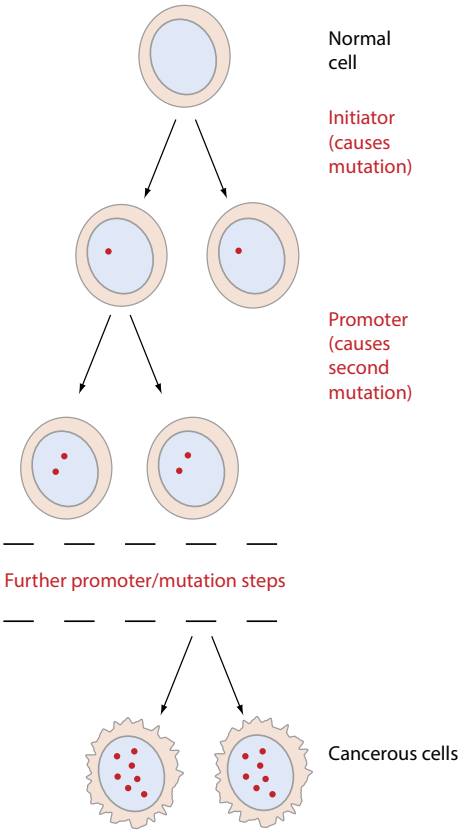


Figure 17.5 Schematic illustrating the multistep theory of cancer.

BOX 17.1 TELOMERES AND TELOMERASE REVERSE TRANSCRIPTASE

Telomeres are highly conserved tandem nucleotide repeats at the ends of chromosomes in eukaryotes (see *Figure 18.10*). They help preserve genomic integrity, protect chromosomes from digestion by exonucleases and prevent aberrant activation of a DNA damage response. Human telomeres are composed of repetitive TTAGGG oligonucleotides and telomere-specific binding protein complexes called telosomes or shelterins. The length of human telomeric DNA is up to 10–15 kb, but each cell division results in the loss of 50–100 bp from telomeres due to a limitation in DNA polymerase activities. Thus, in normal somatic cells, after a set number of mitotic divisions, that is, the Hayflick limit (*Chapter 18*), telomeres become critically short and can no longer be replicated; hence, cells cannot divide, become senescent or die by apoptosis (see *Box 17.3*).

Telomere shortening can be counteracted by the activity of telomerase (see *Figure 18.11)*, which is a ribonucleoprotein complex composed of an RNA-dependent DNA polymerase called telomerase reverse transcriptase (TERT), an RNA subunit called TERC (also known as hTR) and a number of accessory proteins. The catalytic subunit TERT can maintain telomeric lengths by adding telomere repeats at the 3′ end of the lagging strand during DNA replication. The activity of TERT is silenced in most normal adult somatic cells, except in stem cells and some immune cells, for example T cells (*Chapter 4*), which transiently activate telomerase during proliferation. However, the overexpression of *TERT* is especially prevalent in diverse types of tumours. This can be achieved by the amplification of *TERT*, point mutations in relevant promoters or the insertion of viral enhancers upstream of the gene. *TERT* can also be activated by the transcription factors c-*MYC* and NF-κB, and performs functions additional to telomere protection that help cancer progression.

Telomere maintenance via telomerase activation is a nearly universal hallmark of cancer cells that enables them to have unlimited replicative potential. Telomerase can therefore serve as a potential biomarker and as a target in cancer therapy. It has been an attractive target in studies on cancer therapy since its discovery in 1984, but the development of telomerase inhibitors has been challenging and, thus far, no drugs exploiting it have been clinically approved. However, the telomerase inhibitor imetelstat is undergoing clinical trials for treating adult transfusion-dependent anaemia patients with low to intermediate-1 risk of myelodysplastic syndromes (MDSs) who fail to respond to, have lost responses to or are otherwise ineligible for erythropoiesis-stimulating agents.

The massive large-scale genomic DNA analyses required for such studies have only been possible because of the development of next-generation sequencing (NGS) methods (*Box 17.2*), which increased greatly the understanding of the genomic drivers of cancer development and progression.

BOX 17.2 NEXT GENERATION SEQUENCING METHODS

Next-generation sequencing (NGS) methods of sequencing nucleic acids have largely superseded the dideoxy method developed by Fred Sanger in the 1970s, which was used in the Human Genome Project (*Margin Note 7.4*). Several *second*-generation platforms, such as Roche pyrosequencing, Ion Torrent sequencing and Illumina sequencing, are currently available. Each has different methods of sequencing; however, their use has significantly increased the speed of DNA sequencing compared with the Sanger dideoxy method, giving them many applications in genomics research and clinical diagnostics. *Third*-generation sequencing methods include PacBio and Oxford Nanopore sequencing methods. These methods can parallel sequence millions to billions of much larger DNA fragments than earlier methods. The PacBio sequencing method uses fluorescently labelled nucleotides in a single-molecule, real-time (SMRT) approach, and can sequence DNA oligonucleotides up to tens of kilobases long. The Oxford Nanopore sequencing method (*Figure 17.6*) is based (as the name suggests) on nanopore technology. Single-stranded DNA molecules are passed through a nanopore. A potential difference is applied across the pore, and different types of bases produce recognisable changes in the electrical current; hence, the sequences of short to ultralong fragments of DNA are determined in real time as they passes through the pore. Complete DNA sequences are determined using computer algorithms that align the fragments using overlapping sequences.

Next-generation sequencing techniques are applied to determine the sequences of whole genomes (WGs), targeted gene panels, RNAs (mRNA, noncoding RNA and RNA viruses) and the sites of DNA/RNA methylation. These advanced methods are inexpensive and rapid: a human genome can be fully sequenced for several £100s in two to three days and can quickly identify clinically relevant mutations. Thus, NGS has allowed the development of personalised medical therapies (see *Margin Note 17.7*), improved diagnostic methods and evidence-based clinical approaches to treatment regimens.

Figure 17.6 Sequencing nucleic acids using a nanopore system. Courtesy of Oxford Nanopore Technologies (2023), UK.

Genes associated with cancer are often referred to as 'cancer genes' but they are, of course, aberrant, that is, mutated (variant) forms of wild-type genes (*Chapter 15*). Many of them encode molecules that control the cell cycle and regulate cell proliferation and differentiation, while some inhibit apoptosis, a form of genetically programmed cell death (see *Box 17.3*). Mutations in some of these genes may lead to increased proliferation and failures in apoptosis.

Two major groups of genes linked to cancers are:

1. oncogenes
2. tumour suppressor genes

Oncogenes

Oncogenes are mutated forms of wild-type ('normal') genes, called proto-oncogenes, that stimulate an increased proliferation of abnormal cells by encoding proteins such as growth factors and receptors, as well as proteins involved in cell growth and survival, or signal transduction (*Chapter 7*). Examples of oncogenes include *BRAF*, *MYC*, *FOS* and the *RAS* family. Mutation of a proto-oncogene to form an oncogene usually results in the production of a protein that has increased biological activity or in the biosynthesis of excessive amounts of the protein, for example when the gene is continually active. Such mutations are known as 'gain-of-function' mutations.

Oncogenes were first discovered in certain viruses that cause cancers (*Section 17.5*). Such viruses have a gene that is the equivalent of a cellular proto-oncogene. These genes are thought to have been derived initially from host DNA into which the viral DNA was integrated, and to have then undergone mutation during viral passage. Hence, when DNA derived from the virus becomes inserted into the host genome, the viral oncogene may provide the additional signal necessary for cellular proliferation, or it may override the normal cellular controls and result in the unregulated division of infected cells. The viral and cellular forms of the oncogene are usually distinguished by the prefixes '*v*' and '*c*', respectively, for example v-*MYC* and c-*MYC*. Examples of oncogenes and associated tumours are shown in *Table 17.3*.

Margin Note 17.3 The oncogene *PIK3CA* and breast cancer

The oncogene *PIK3CA* located at 3q26.32 encodes the p110α catalytic subunit of phosphatidylinositol-3-kinase (PI3K), which *in vivo* catalyses the phosphorylation of phosphatidylinositol 4,5-bisphosphate (PIP_2) to phosphatidylinositol 3,4,5-triphosphate (PIP_3). This reaction is part of the PI3K pathway, which regulates cellular proliferation, survival and metabolism. Constitutive activation of the PI3K pathway is commonly related to tumorigenesis. Mutations in *PIK3CA* have been reported in 20–40% of breast cancer patients. An inhibitor of *PI3KCA*, alpelisib, in combination with fulvestrant is approved for the clinical treatment of *PIK3CA* mutated, HR[a] positive, HER2[b] negative advanced breast cancer.

[a] HR, receptors for the hormones oestrogen and progesterone; [b] HER2, epidermal growth factor receptor 2.

Proto-oncogene (site)	Gene product	Examples of associated cancers
BCR-ABL (t(8;22)(p11;q11))	Tyrosine kinase	Chronic myelogenous and acute lymphoblastic leukaemias
BRAF (7q34)	Serine/threonine kinase	Colorectal, non-small cell lung and thyroid cancers; melanoma
ERBB1/HER1/EGFR (7p11.2)	Epidermal growth factor receptor (EGFR)	Non-small cell lung cancer; glioma
ERBB2 (HER2, NEU) (17q12)	Human epidermal growth factor receptor 2	Breast, gastric and ovarian cancers; glioblastoma
FOS/AP-1 (14q24.3)	Transcription factor	Osteoblastoma
HRAS (11p15.5)	GTPase	Bladder and thyroid cancers; rhabdomyosarcoma
KRAS (12p12.1)	GTPase	Colorectal, gastric, lung, pancreatic and thyroid cancers; acute myeloid leukaemia
MYC (8q24.21)	Transcription factor	Burkitt lymphoma; multiple myeloma
PIK3CA (*Margin Note 17.3*) (3q26.32)	Catalytic subunit of phosphatidylinositol-3-kinase	HR[a] positive, HER2[b] negative breast cancer
SRC (20q11.23)	Tyrosine kinase	Colon cancer and endometrial carcinoma
TERT (*Box 17.1*) (5p15.33)	Telomerase reverse transcriptase	Hepatocellular carcinoma; glioblastoma; melanoma, skin basal cell and skin squamous cell carcinomas; medulloblastoma

[a] HR, receptors for oestrogen and progesterone; [b] HER2, epidermal growth factor receptor 2.

Table 17.3 Examples of oncogenes and their associated tumours. ⏎

Tumour suppressor genes

Tumour suppressor genes, in contrast with proto-oncogenes, encode proteins that inhibit the proliferation of cells that contain deleterious mutations. Some tumour suppressor proteins have crucial functions in inducing apoptosis (see *Box 17.3*) or in the repair of damaged DNA. Thus, loss of function mutations in tumour suppressor genes may then lead to dysregulated cell growth. Examples of tumour suppressor genes are those for the:

- tumour protein 53 (*TP53*)
- retinoblastoma susceptibility gene (*RB1*)
- breast cancer susceptibility genes 1 and 2 (*BRCA1 and BRCA2*)

Table 17.4 lists other examples.

Tumour protein 53 gene (TP53)

The gene *TP53* is found on the short arm of chromosome 17. It encodes a phosphoprotein, called tumour suppressor protein p53, which has been called the 'guardian of the genome'. The concentration and activity of p53 are low in non-stressed conditions mainly through its degradation by the E3 ubiquitin ligase mouse double minute 2 (MDM2)-proteasome system (*Margin Note 17.4*). However, cellular stresses, such as DNA damage, oncogene activation, hypoxia or replication stress, can all result in p53 post-translational modifications that lead to its activation. Active p53 can bind to target DNA (*Figure 17.7*), blocking the division of damaged cells by arresting the cell cycle at G1 phase (*Chapter 15*), initiating DNA repair and inducing senescence or apoptosis, thus preventing the replication of potential tumour cells. In addition to its function as a transcription factor, p53 exerts tumour-suppressive actions by interacting with other proteins, for example BCL-2, BCL-xL and breast cancer type 1 susceptibility protein (BRCA1). Mutations in *TP53* are found in over 50% of human cancers and across a variety of cancer types and can lead to the production of defective p53 that is unable to recognise its DNA binding sites. Thus, the replication of the cell is not inhibited, leading to a failure in removing damaged and potentially malignant cells.

Retinoblastoma susceptibility gene (RB1)

The tumour suppressor gene *RB1* is located at 13q14.2. It encodes a nuclear phosphoprotein, pRB. The protein is a negative regulator of the cell cycle, whose activity depends on its level of phosphorylation. Hypophosphorylated pRB arrests the cell cycle at the G1

Margin Note 17.4 ⏎
E3 ubiquitin-protein ligase or mouse double minute 2

The concentration and activity of p53 is low in non-stressed cells, mainly because of its degradation in multisubunit structures called **proteasomes**. Proteins destined for degradation in this manner are tagged with the polypeptide ubiquitin (M_r 8500) in an ATP-dependent manner, which requires the sequential activities of three enzymes (E1, E2 and E3). The third, E3, is a ubiquitin-protein ligase that attaches ubiquitin to lysine residues of the protein allowing its recognition and proteolysis by proteasomes. The ubiquitin is cleaved off and recycled during proteolysis. In humans, the E3 active in the proteasome system responsible for degrading p53 is the enzyme mouse double minute 2 (MDM2), also called E3 ubiquitin-protein ligase MDM2, which is encoded by *MDM2* located at 12q15. In addition to its role as a negative regulator of p53, MDM2 also inhibits transcription of *TP53*. The oncoprotein form of MDM2 therefore promotes cell survival, proliferation, invasion and therapeutic resistance in many types of cancer. Numerous inhibitors of MDM2, for example nutilins like *cis* imidazoline analogues, are under clinical development for possible use as cancer treatments.

Tumour suppressor gene (site)	Gene product	Examples of associated cancers
APC (5q22.2)	A regulator of the WNT[a] signalling pathway	Colorectal, pancreatic and prostate cancers; familial adenomatous polyposis
ATM (11q22.3)	Serine/threonine kinase	Glioma, leukaemia, lymphoma and medulloblastoma
BRCA1 (17q21.31)	DNA repair-associated protein	Breast and ovarian cancers
BRCA2 (13q13.1)	DNA repair-associated protein	Breast, ovarian and pancreatic cancers
CDKN2a (9p21.3)	Cyclin-dependent kinase inhibitor	Melanoma, pancreatic cancer
KMT2C (7q36.1)	Methyltransferase	Breast, lung and colorectal cancers
MLH1 (3p22.2)	DNA mismatch repair protein	Lynch syndrome; colorectal, endometrial and ovarian cancers
MSH2 (2p21)	DNA mismatch repair protein	Colorectal and ovarian cancers, hereditary nonpolyposis colon cancer (HNPCC); endometrial carcinoma
RB1 (13q14.2)	Nuclear phosphoprotein	Small cell lung cancer, osteosarcoma and retinoblastoma
TP53 (17p13.1)	Transcription factor	Glioma; Li–Fraumeni syndrome; breast, colorectal and lung cancers, sarcoma
VHL (3p25.3)	A component of ubiquitin ligase	Renal cell carcinoma; Von Hippel–Lindau syndrome

[a] WNT, wingless-related integration site, is a family of proteins that play key roles in regulating cell-to-cell communication during embryogenesis, cell proliferation, differentiation, migration and stem cell renewal.

Table 17.4 Examples of tumour suppressor genes and their associated tumours.

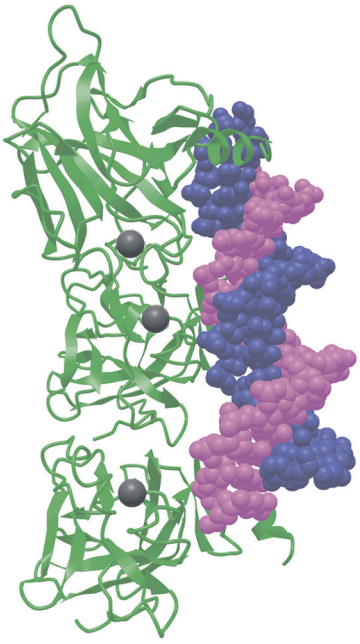


Figure 17.7 Molecular model showing the binding of p53 to DNA (blue and magenta). This inhibits the division of damaged cells that are potentially cancerous. PDB file 1TUP.

BOX 17.3 APOPTOSIS AND CANCER

The number of cells in the body is tightly regulated to achieve a balance between those formed by mitosis and those that die by accidental death or by regulated/programmed cell death pathways, such as apoptosis, necrosis and autophagy. The most common form of genetically programmed death is **apoptosis** (from the Greek *apo* and *ptosis*, meaning away and falling, respectively), which accounts for the deaths of 10^{10}–10^{11} cells daily. In normal cells, apoptosis is held in balance by proapoptotic and antiapoptotic proteins. Members of the B cell lymphoma 2 (BCL2) protein family are key regulators of apoptosis including antiapoptotic proteins such as BCL2, which gives the group its name, and BCL-xL, or proapoptotic proteins, such as PUMA, BAX, BAD and BID.

Apoptosis can occur naturally to remove, for example, superfluous cells during development. However, it can also be induced when a cell is subjected to stresses by irreparable damage to its DNA (mutations), oxidative stress or radiation or is infected with a virus, which stimulates the *intrinsic* apoptotic pathway. In such cases, the transcription factor p53 can initiate the biosynthesis of proteins, such as PUMA (p53 upregulated modulator of apoptosis), which promotes apoptosis by binding to antiapoptotic proteins. Proapoptotic proteins form multisubunit channels through the outer mitochondrial membrane allowing cytochrome *c* to be released into the cytosol. Within the cytosol, cytochrome *c* combines with factors such as apoptotic protease-activating factors (Apaf-1) and the protease procaspase-9 to form a multimeric complex called an **apoptosome** (*Figure 17.8*). Procaspase-9 is a so-called initiator caspase. In the apoptosome, it is converted to the activate caspase-9, which catalyses the effector or executioner caspases, such as caspase-3, -6 and -7, which initiate cell degradation. Following activation, effector caspases hydrolyse substrates and activate other proteinases in a proteolytic cascade that amplifies the proteolytic response to the original apoptotic signal. (The use of cascade mechanisms to amplify a response is a common feature in biology. Examples include the activation of complement and signal transduction mechanisms [*Chapters 4* and *7*, respectively].) Substrates of effector caspases are varied and include 1000s of proteins, for example structural nuclear lamins and cytoskeleton proteins in the cytoplasm, as well as intercellular binding integrins, a protein inhibitor of the caspase-activated DNase (CAD) and proteins with roles in repairing DNA and regulating the cell cycle (*Chapter 15*). During the early stages of apoptotic death,

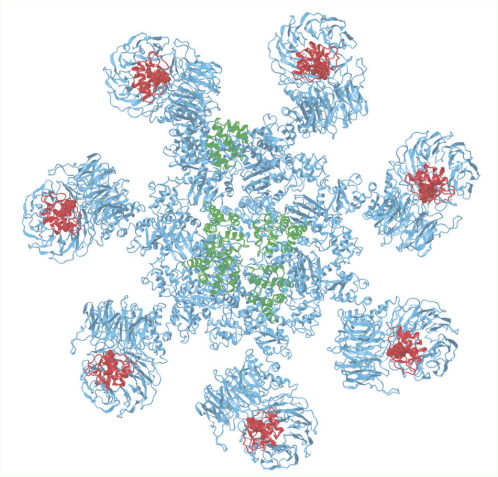


Figure 17.8 Molecular model of an apoptosome. Caspases are shown in green, apoptotic protease-activating factors (Apaf-1) in blue and cytochrome *c* in red. PDB file 5WVE.

the cell shrinks, phosphatidylserine is exposed on the cell surface and cytoplasmic vesicles fuse with the plasma membrane forming bubble-like extensions or blebs on the cell surface. Chromatin condenses and the chromosomal DNA is fragmented. The contents of apoptotic cells are finally split into distinct membrane-enclosed vesicles, called apoptotic bodies. Apoptotic bodies are rapidly engulfed by phagocytes, for example macrophages, where, eventually, the cell remnants are degraded. Thus the contents of apoptotic cells are retained within the plasma membrane and are not released to stimulate inflammatory responses (see *Box 4.1*), which could damage neighbouring cells. Additionally, the secretion of cytokines such as interleukin 10 (IL-10) and transforming growth factor β (TGF-β) by the phagocytic cells inhibits inflammation (*Chapter 4*).

Cells that express integral plasma membrane proteins called death receptors (DRs) can initiate an *extrinsic* apoptotic pathway when they receive appropriate signals from another cell. Cells destined for apoptosis express DRs, such as Fas and the receptor for tumour necrosis factor-α (TNF-α; see *Figure 17.24*) on their surfaces. This ensures they are targets of killer cells, such as cytotoxic T lymphocytes (see *Figure 6.39*) and natural killer cells (*Chapter 4* and *Figure 4.17*), which express Fas ligands (FasLs) on their surfaces. For example, when FasLs bind to three Fas receptors on the target cell, they promote a conformational change in Fas that allows the adaptor protein, Fas-associated death domain protein (FADD), to be recruited to the receptors. The resulting complex is then able to recruit procaspases-8 and -10 to form a death-inducing signal complex (DISC). The low latent protease activities of the procaspases can now mutually activate each other forming an active initiator caspase-8/-10 that can catalyse the conversion of executioner procaspases to active executioners, and the cell then enters apoptosis, with its degradation occurring as described above. Additionally, p53 upregulates Fas, meaning the intrinsic and extrinsic pathways can be linked to amplify the death response. Soluble TNF-α is produced by immune system cells and binds to the DR TNF receptor 1 on the surface of cells programmed for apoptosis. Binding stimulates the formation of DISC and the executioner caspase cascade.

Defects in the genes controlling apoptosis are involved in many clinical conditions. These include Alzheimer and Parkinson diseases (*Chapter 18*) and autoimmune disorders, for example rheumatoid arthritis, where healthy cells die (*Chapter 5*). However, in neoplasia or cancer development, cells that should be eliminated are protected from apoptosis. Evasion of apoptosis is one of the hallmarks of cancer cells. Since the discovery of apoptosis, following studies by Sydney Brenner (1927–2019), John Sulston (1942–2018) and H. Robert Horvitz (1947–), who shared a Nobel Prize in 2002, a goal of clinical oncology has been to develop effective ways of stimulating apoptosis in cancer cells as a means of eliminating tumours. To date, numerous low M_r agents that target BCL-2 members are subject to various stages of clinical trials; one, venetoclax, has been approved for anticancer therapy. Venetoclax is used to treat adults with chronic lymphocytic leukaemia, small lymphocytic lymphoma or acute myeloid leukaemia (*Section 17.8*). Venetoclax binds to BCL-2 (*Figure 17.9*) and selectively inhibits its actions. This alters the cellular balance in favour of apoptosis, leading to the death of the cancer cell and slowing the progression of the disease. Death receptors/the extrinsic apoptotic pathway are also targets in cancer therapy. For example, established chemotherapeutic agents that damage DNA and target Fas expression include cisplatin and methotrexate (see *Figures 17.36* and *17.32*, respectively). A number of potential therapeutic agents have also been proposed to stimulate the apoptotic function of DRs and ligands of the extrinsic pathway, such as inhibitors of histone deacetylase (HDAC), proteasome inhibitors, cyclooxygenase-2 inhibitors and several antibodies that specifically target DRs.

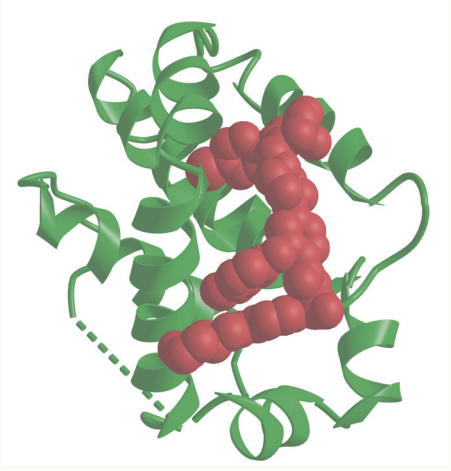


Figure 17.9 Molecular model of venetoclax, shown in red, bound to BCL-2, depicted in green. PDB file 6O0K.

restriction point by binding E2F transcription factor, which is essential for the expression of genes involved in the cell cycle G1/S phase transition (*Chapter 15*). The pRB also stabilises constitutive heterochromatin to maintain the overall chromatin structure. Inactivation of pRB can result in the impairment of the cell cycle and uncontrolled cell proliferation. Mutations in *RB1* are associated with retinoblastoma, a malignant tumour in the retina (see *Boxes 17.5* and *17.9*).

Breast cancer susceptibility genes 1 and 2

Mutations in *BRCA1* and *BRCA2*, which are associated with an increased susceptibility to cancer of the breast and ovaries, account for less than 2% of all breast cancers. However, patients with defective *BRCA* variants have a much greater risk of developing breast cancer than those who do not. More than 1600 mutations have been identified in *BRCA1* (sited at 17q21), which include deletions, insertions and many single-nucleotide polymorphisms in both its introns and exons (see *Margin Note 13.7*). Women who inherit a mutated *BRCA1* have a 60–83% chance of contracting breast cancer and a 20–40% chance of developing ovarian cancer. The gene encodes BRCA1, a complex protein composed of several domains (*Figure 17.10*) that are essential for its multiple functions. These include regulating the transcriptional activity of p53 and ubiquitin protein ligase activity.

Both BRCA1 and BRCA2 are involved in repairing double-stranded breaks (DSBs) in DNA by regulating homologous recombination (HR). Repair can occur in the S and G2 phases of the cell cycle when sister chromatids are available as templates for HR. During HR repair, BRCA1 is necessary for the recruitment of BRCA2 to the sites of DNA DSBs. It also regulates nonhomologous end-joining (NHEJ) repair, which is one of the main DNA repair pathways when the broken DNA ends are directly ligated when an homologous DNA template is not available. Homologous recombination is essentially an error-free process, but it requires an extended period for completion. In contrast, NHEJ is an error-prone DNA reparation process that causes mutations at the site of damage. The dephosphorylation of p53-binding protein 1 by BRCA1 promotes HR-dependent DNA repair, switching the repair pathway from NHEJ to HR.

Given both BRCA1 and BRCA2 play crucial roles in the maintenance of genomic integrity, the presence of a mutated *BRCA1*, with the production of a dysfunctional BRCA1, leads to a failure in eliminating damaged cells. Women carrying mutated forms of *BRCA1* and having a strong family history of breast cancer may be offered a **prophylactic bilateral mastectomy**, that is, the removal of the apparently healthy breasts to prevent any development of breast cancer. Mutations in *BRCA2* sited at 13q13.1 confer a 40–60% and 10–20% chance of developing breast cancer and ovarian cancer, respectively. Mutations in *BRCA2* also increase the risk of male breast cancer.

The gene *CHEK2* (formerly *CHK2*), located at 22q12.1, encodes a checkpoint kinase 2 (CHK2), which is involved in the control of the cell cycle, apoptosis, senescence and DNA repair and is also associated with breast cancer. The enzyme normally occurs as an inactive monomer. Upon DNA damage such as DSBs, ataxia telangiectasia-mutated (ATM) kinase phosphorylates CHK2, which promotes its homodimerisation to an active kinase. Active CHK2 phosphorylates many intracellular protein targets, including p53, which increases its stability and activity. Active CHK2 also phosphorylates BRCA1, which is believed to modulate its function in HR repair. Multiple variants of *CHEK2* have been described; inheriting pathogenic types increases the risk of both women and men developing breast cancer.

17.5 CAUSES OF CANCER

Mutations that lead to cancerous states are caused by or associated with a number of factors. These include errors in replicating DNA or failures in repairing damaged DNA, or they may be induced by a variety of environmental agents, including chemical carcinogens and ionising radiation, or by the actions of some viruses. Mutations in cancer-associated germ line genes may also mean that offspring carrying the variant gene are more susceptible to developing cancer.

Inherited cancers

The association between the inheritance of a mutated gene and the development of cancer may be direct or indirect. For a direct association, the offspring must inherit a mutated gene that confers increased susceptibility to a specific type of cancer. Examples of cancer-associated genes with direct effects include both *BRCA1* and *BRCA2*, as described in *Section 17.4*. In indirect associations, the inherited gene is linked to defective mechanisms for repairing DNA that, in turn, lead to a greater likelihood of cancer. An example of the latter occurs in patients with the disease xeroderma pigmentosum in whom a failure to repair ultraviolet (UV) light-induced mutations in the DNA of skin cells leads to the development of multiple skin cancers.

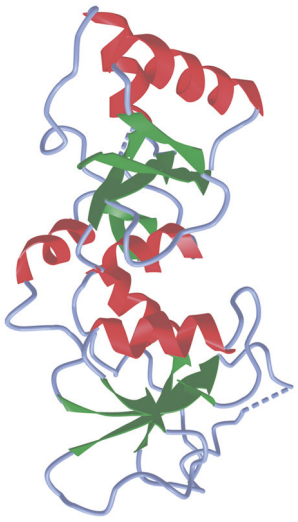


Figure 17.10 Molecular model of a BRCA1 protein. PDB file 1L0B.

Margin Note 17.5
Knudson, the two-hit theory and *RB1*

Alfred Knudson (1922–2016) found that patients with familial retinoblastomas presented at an early onset age and developed tumours in both eyes, while patients in nonfamilial cases developed a unilateral tumour later in childhood. In 1971, Knudson proposed the 'two-hit' theory regarding the genetic basis for retinoblastoma, suggesting that two mutations were required to initiate the development of retinoblastoma (see *Box 17.9*) and that the gene involved likely acted in a recessive manner. This theory was used to explain both heritable and spontaneous cases of retinoblastoma: the former had one hit in a tumour suppressor gene in the germ line and only required one more hit in a somatic retinal cell, whereas the latter required the first and second mutation to occur in a somatic cell. This model explained why spontaneous or nonfamilial cases of retinoblastoma occurred later in life and were never bilateral. Remarkably, Knudson described the existence and properties of a retinoblastoma tumour suppressor gene, *RB1*, 15 years before the gene was cloned and even longer before the biochemical roles of pRB in regulating cell proliferation were understood in any detail.

Today, this 'two-hit' theory can be modified to encompass both genetic and epigenetic inactivation of tumour suppressor genes. It still serves as a basis for the understanding of cancer genetics.

BOX 17.4 LI FRAUMENI SYNDROME

Li–Fraumeni syndrome (LFS) is a rare autosomal dominant syndrome that is inherited in typical Mendelian fashion (*Figure 17.11* and *Chapter 15*). The syndrome has been linked to mutations in *TP53* and predisposes the patient to developing cancer. Indeed, approximately 70% of members of LFS families have *TP53* variants. Mutations can occur spontaneously in germ cells of one parent or they may occur early in the development of the embryo. Some LFS families have an oncogenic variant of *CHEK2* rather than a mutated *TP53*.

Li–Fraumeni syndrome was first described in 1968, and more than 500 families had been reported by 2023. The lifetime risk of cancer in individuals with LFS is more than 70% and 90% for men and women, respectively. Within LFS families, affected individuals develop cancer at an early age. Family members show a wide range of tumours, including cancers of the breast, brain, soft tissue, bone and adrenal cortex, as well as leukaemias, which are diagnostic for the syndrome. However, patients within LFS families have also been found to develop tumours at other sites in addition to those used in diagnosis. Children who survive an initial cancer have a high risk of developing a second.

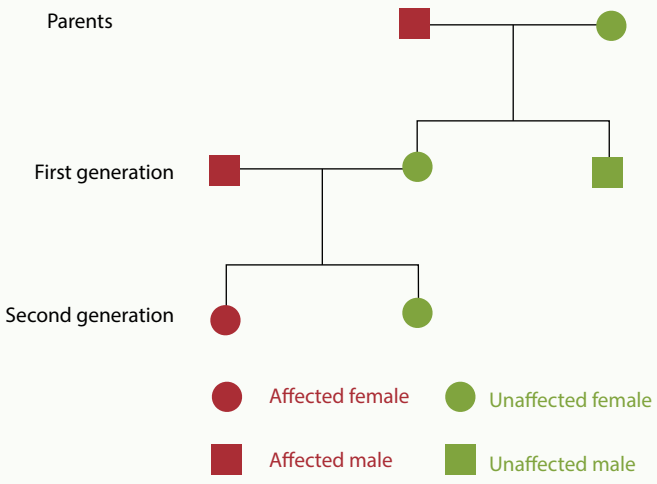


Figure 17.11 Scheme showing the autosomal dominant mode of inheritance of Li–Fraumeni syndrome. ⏎

To diagnose familial LFS, patients must meet all three classic clinical criteria: have a sarcoma diagnosed before the age of 45 years, a first-degree relative (parent, sibling or child) diagnosed with any cancer before the age of 45 years, or a first-degree or second-degree relative (grandparent, aunt/uncle, niece/nephew or grandchild) diagnosed with any cancer before the age 45 years or a sarcoma diagnosed at any age. The age at which patients present with an initial primary tumour is significant, since more than half of patients with LFS present before the age of 45, compared with 10% of the general population.

The most common cancer in females with LFS is breast cancer. Regular **mammography**, which is an X-ray examination of the breast, will detect a tumour as a shadow on the X-ray film or 'breast mammogram' and has been advocated for women in affected families. However, regular X-ray screening is controversial, since ionising radiation itself may increase the risk of cancer in these patients. Prophylactic bilateral mastectomy has also been proposed, although, again, this has aroused debate, since the women remain at risk of developing a cancer at a different location. Genetic counselling of individuals in affected families is essential to inform patients of the potential risks of cancer and to help them develop strategies to avoid at-risk behaviours, such as exposure to radiation.

Chromosomes and cancer

Chromosomal abnormalities are a common characteristic of cancer. The numerical and structural chromosome abnormalities found in cancers include aneuploidy, chromosomal deletions, amplifications or translocations that result in genomic instability (*Chapter 15*). For example, about 95% of patients with chronic myelogenous leukaemia (CML; *Section 17.8*) have a translocation involving chromosomes 9 and 22, that is, t(9:22). The translocation produces a longer chromosome 9 and a shorter chromosome 22, commonly called the Philadelphia or Ph chromosome (*Figure 17.12*). The Ph chromosome is also found in 25–30% of adults with acute lymphoblastic leukaemia (ALL; *Section 17.8*). The translocation results in the fusion of the *BCR* gene on chromosome 22 and part of the *ABL* gene on chromosome 9 (*Figure 17.13*). The fused *BCR–ABL* encodes a tyrosine kinase that activates multiple intracellular signalling pathways. These include the RAS/RAF/MAPK, PI3K/AKT/mTOR and WNT/β-catenin pathways, which are involved in regulating cell proliferation, differentiation, survival and migration. In healthy cells, these pathways would be activated by, for example, the binding of growth factors to their cellular receptors. However, the continuous expression of BCR–ABL constantly stimulates such pathways, which, in turn, promote the proliferation and the development of CML. In patients, the translocation originally occurred in a single bone marrow cell. However, clonal expansion of the cell results in the blood becoming populated with cells bearing the Ph chromosome. During active disease, additional chromosomal abnormalities appear and these are indicative of a poor prognosis. Several drugs, such as the tyrosine kinase inhibitors imatinib and nilotinib (*Figure 17.14*), which target BCR–ABL by blocking its binding to ATP, are used to treat CML.

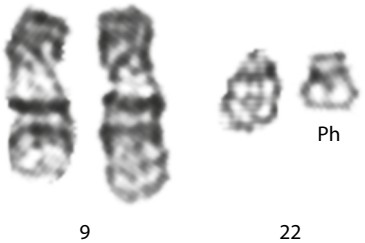


Figure 17.12 The abnormally long chromosome 9 and the Philadelphia chromosome (Ph, abnormal chromosome 22) from a patient with chronic myelogenous leukaemia (CML). Reprinted with permission of the Wisconsin State Laboratory of Hygiene, University of Wisconsin Board of Regents. ⏎

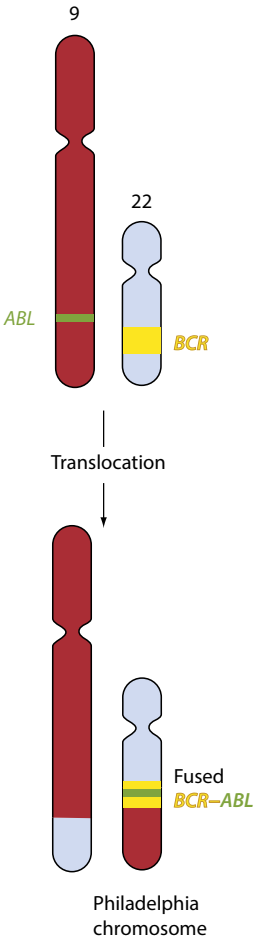


Figure 17.13 Schematic showing the translocation between chromosomes 9 and 22, which results in the Philadelphia chromosome.

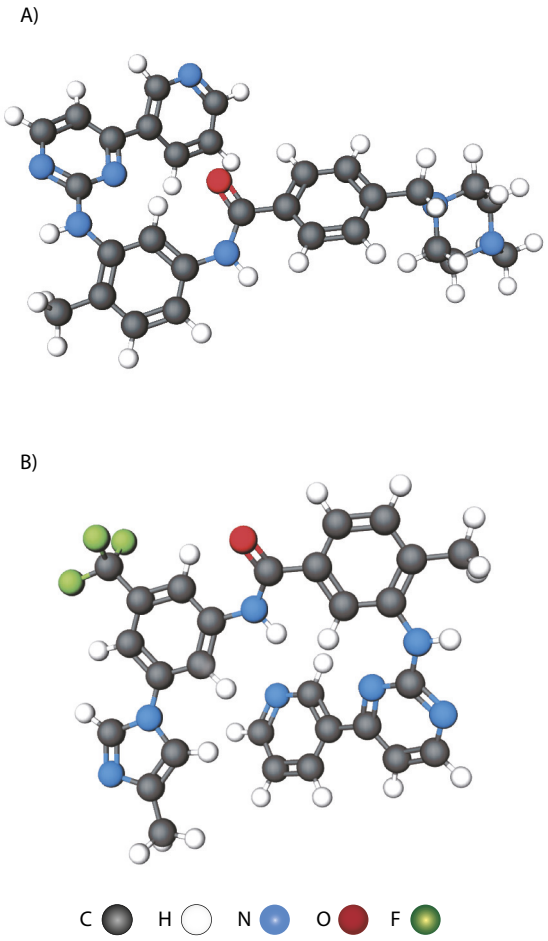


Figure 17.14 Computer drawn models of (A) imatinib and (B) nilotinib.

Other chromosome abnormalities have been associated with a wide range of cancers, including breast cancer, prostate cancer and neuroblastoma, but none to the same extent as that found in CML, where the presence of the Ph chromosome is diagnostic for the disease. Other examples of chromosome abnormalities associated with cancer are shown in *Table 17.5*.

Abnormality	Associated cancer(s)
Amplification on the long arm of chromosome 17 (amp(17q))	Breast carcinoma
Deletion in the short arm of chromosome 1 (del(1p))	Colorectal adenoma
Deletion in the short arm of chromosome 3 (del(3p))	Non-small cell carcinoma of the lung; renal cell carcinoma
Deletion of band 13 in the short arm of chromosome 11, that is, del(11p13)	Wilms tumours
Translocation between chromosome 1, band 36 on the short arm, and chromosome 17, band 12 on the long arm (t(1;17)(p36;q12))	Neuroblastoma
Translocation between chromosome 2, band 37 on the long arm, and chromosome 13, band 14 on the long arm (t(2;13)(q37;q14))	Rhabdomyosarcoma (alveolar)
Translocation between chromosome 15, band 22 on the long arm, and chromosome 17, band 21 on the long arm (t(15;17)(q22;q21))	Acute promyelocytic leukaemia (*Margin Note 17.18*)

Table 17.5 Examples of chromosomal abnormalities found in different types of cancers.

Chemical carcinogens

In 1775, Percivall Potts (1714–1788), an English surgeon, noted the high incidence of scrotal cancer amongst chimney sweeps in London and suggested that this may be related to the accumulation of soot in their clothes. As a result of legislation introduced to ensure that chimney sweeps were able to bathe and to change their clothing regularly, scrotal cancer was eliminated in this profession. In 1918, Katsusaburo Yamagiwa (1863–1930) and Koichi Ichikawa (1888–1948) showed that tumours could be induced experimentally by painting coal tar on the skin of rabbits. This property of certain compounds to induce tumours experimentally (*Figure 17.15*) led to the identification of many carcinogenic chemicals, including those listed in *Table 17.6*.

Figure 17.15 Carcinogen-induced tumours on the skin of a rat. ⏎

Chemical	Source	Cancers caused
Aflatoxin	*Aspergillus flavus* and *A. parasiticum* growing on crops such as corn and peanuts	Liver cancer
Arsenic	In insecticides and herbicides	Lung and skin cancers
Asbestos	Mineral fibres used in fire-insulation and brake linings (produced before 1990)	Mesothelioma
Benzene	Petrochemicals, dyes (industrial exposure)	Bladder cancer
Benzo[a]pyrene	Partial combustion of petroleum; tobacco smoke	Lung
Diethylstilbestrol	Once used to prevent miscarriage in women	Vaginal tumours in offspring of treated women
Polychlorinated biphenols	Industrial processes, insecticides	Liver and skin cancers
Polycyclic aromatic hydrocarbons	Partial combustion of petroleum	Lung cancer
Vinyl chloride	Industrial processes	Liver cancer

Table 17.6 Examples of carcinogenic chemicals listed by the International Agency for Research on Cancer (IARC) and the USA National Toxicology Program (NTP). ⏎

The chemical induction of cancers is assumed to be a multistep process, probably involving mutations in several genes, possibly on different chromosomes (see *Figure 17.5*). Carcinogenic compounds mutate DNA usually by one of a limited number of reactions. For example, alkylating agents such as dialkylnitrosamines and aflatoxin B1 lead to the addition of alkyl groups to electron-rich sites in DNA (see *Figure 12.8*). Aromatic amines and amides form highly electrophilic aryl nitrenium ions, which also interact with DNA. Polycyclic aromatic hydrocarbons (PAHs) are a large group of organic compounds with two or more fused aromatic rings. They have a relatively low solubility in water, but are highly lipophilic and so can be transported by lipoproteins in the blood and pass through biological membranes. On entering cells of the lungs, GIT and/or skin, they can form DNA-adduct products and also induce oxidative stress, both of which can induce mutations.

The process of chemical carcinogenesis results from the actions of combined chemical, physical, biological and genetic insults. At its simplest, it can be thought to occur in three phases:

1. initiation
2. promotion
3. progression

During tumour initiation, the carcinogen, or a metabolite of the carcinogen, produces a mutation. The cell may repair the damage but misrepair may lead to heritable changes. A cell that has undergone initiation, however, is not yet cancerous because the cell has to 'escape' from normal growth control and become autonomous when tumour promoters stimulate the clonal proliferation of an initiated cell (*Figure 17.5*). Tumour progression involves the additional changes leading to malignancy and metastasis. Emerging evidence suggests that tumour microenvironment (TME) plays critical roles in tumour growth, progression, metastasis and therapeutic responses (*Box 17.5*).

BOX 17.5 TUMOUR MICROENVIRONMENT

The tumour microenvironment (TME) consists not only of neoplastic cells but also of stromal cells and non-cellular components. The non-cellular components of extracellular matrix (ECM), such as collagens, fibronectin, hyaluronans and laminins, provide architectural scaffolding, mechanical support and anchorage sites for cell adhesion; store water and various growth factors; and induce intracellular signalling pathways via cell–ECM interactions. Due to tumour cell–ECM interactions, tumour cells undergo an epithelial–mesenchymal transition (EMT) and become invasive.

The non-malignant stromal cells include vascular cells (for example endothelial cells, pericytes, fibroblasts and adipocytes) and immune cells (such as lymphocytes, monocytes, tumour-associated macrophages and neutrophils). These cells communicate not only with each other but also with the tumour cells by secreting a variety of molecules that affect tumour behaviour through different intracellular signalling pathways. Several growth factors, such as epidermal growth factor (EGF), fibroblast growth factor (FGF), platelet-derived growth factor (PDGF), transforming growth factor (TGF) and vascular endothelial growth factor (VEGF), have been widely recognised as master regulators of the interactions occurring within the TME. For example, endothelial cells are activated by the VEGF-stimulated signalling pathway, leading to an expansion of the vascular network supplying the tumour, a process referred to as **angiogenesis**, which increases the supply of oxygen and nutrients that the tumour needs for growth.

Cytokines, chemokines (*Chapter 4*), extracellular vesicles, proteolytic enzymes (such as matrix metalloproteinases (MMPs) that remodel the ECM) and microRNAs (miRNAs)[a] in the TME also play pivotal roles in controlling growth, invasion and metastasis of cancer cells. Chemical carcinogens, including asbestos and crystalline silica, are implicated in the tissue remodelling and inflammation (see *Box 4.1*) that accompany cancer development, particularly in lung cancers. Given the increased understanding of the crucial roles of the TME in tumour development and therapeutic resistance, many efforts are being devoted to targeting components of the TME to achieve therapeutic benefits in cancer patients.

[a] microRNAs (miRNAs) are known to interfere with DNA transcription.

Many targets of chemical carcinogens are the proto-oncogenes and tumour suppressor genes described in *Section 17.4*. Mutations in *TP53* have been found in various types of human tumours. The mutations occur at vulnerable sites known as 'hotspots', although the position of the hotspot is not the same for all carcinogens. Approximately 80% of *TP53* mutations are missense mutations, which are mainly located in exons 5–8 that encode its DNA binding domain. For example, the metabolite of benzo[a]pyrene preferentially forms adducts with guanine bases (see *Figure 12.8*) at codons 157, 248 and 273 of *TP53*, which are the same mutational hotspots found in human lung cancers. This, therefore, supports the link between tobacco abuse and lung cancer that was established in the 1950s (*Box 17.6*). Some chemical carcinogens may work by promoting the formation of reactive oxygen species (ROS; see *Chapters 12* and *18*), which also attack DNA. (Thus, health food shops promote the sale of supplements that are known to remove or 'scavenge' ROS.) Alternatively, some carcinogens may interfere more directly with the regulation of cell proliferation or receptor-mediated cell signalling.

BOX 17.6 CIGARETTES AND CANCER

The link between cigarette smoking and lung cancer was first established by the epidemiologist William Doll (1912–2005). Doll qualified in medicine in 1937 and, after World War II, began work on a project to determine the causes of a sharp increase in the number of deaths from lung cancer over the previous 30 years. Originally, he thought the increase might be due to exhaust fumes from cars. His study involved giving questionnaires to lung cancer patients about their habits. From a relatively small-scale study to one which involved questionnaires sent to over 60 000 doctors, he and his colleague, Austin Hill (1897–1991), were able to show that the risk of lung cancer was proportional to the number of cigarettes smoked by the patient (*Figure 17.16*) and that long-time smokers had three times the mortality rate of nonsmokers. In addition, they established the link between cigarette smoking and a number of other serious diseases, such as coronary thrombosis (*Chapter 14*) and chronic bronchitis (*Chapter 5*).

In 2019, there were an estimated 2.5 million cancer-related deaths attributable to smoking globally, which accounts for about 30% of all deaths from cancers. In addition to being an

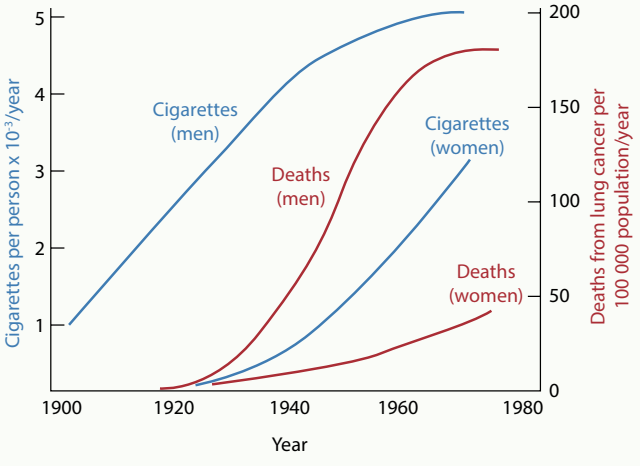


Figure 17.16 The relationship between deaths from lung cancer and the number of cigarettes smoked. Note the 20-year lag between the increases in cigarettes smoked and the increasing numbers of deaths.

established cause of lung cancer, smoking also increases the risk of developing cancers of the bladder, colon, larynx, mouth and oesophagus, liver, kidney, stomach, pancreas and rectum and of leukaemia. Fortunately, knowledge of the link between smoking and cancer, mass media interventions, health warnings on packaging, increasing taxes on tobacco and cigarettes, and providing smoke-free environments have led to a decline in smoking rates in the UK (*Figure 17.17*) and other developed countries.

Doll used epidemiological studies to establish other links between social habits and clinical conditions. For example, he showed that imbibing alcohol during pregnancy can have undesirable effects on the unborn baby and that exposure to relatively small amounts of ionising radiation increases the risk of leukaemia.

Since electronic (e-)cigarettes came onto the market from about 2007, the prevalence of e-cigarettes has increased, particular among young adults. In 2023, it was estimated that the global number of e-cigarette users ('vapers') was approximately 114 million. Electronic cigarettes produce an aerosol by heating a liquid that usually contains nicotine and flavourings, such as diacetyl, and are often packaged in a manner that appeals to younger users. The widespread use of e-cigarettes is still relatively new. Their long-term health effects and carcinogenic risk will need continuous investigation.

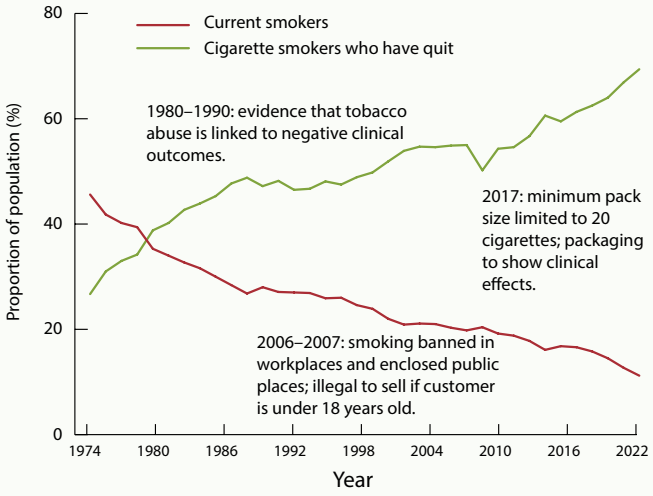


Figure 17.17 The proportion of current smokers and those who had smoked but have now ceased, aged 16 years and over during 1974–2022 in Great Britain. Data from the Office for National Statistics: adult smoking habits in the UK: 2022.

Testing for potential carcinogens

The main test of the ability of chemicals to cause mutations in DNA is the Ames test, which uses a mutant of *Salmonella typhimurium* that is unable to grow on growth media lacking the amino acid histidine (*Figure 17.18*). The test involves exposing the bacterium to the chemical in question, usually in the presence of an extract of mammalian liver to provide enzymes that may be necessary to activate any procarcinogens present. Reversion of the bacterium to one that can biosynthesise histidine and, thus, grow and form colonies on the histidine-deficient medium indicates that the potential carcinogen has mutated its DNA. Potential mutagenicity can also be determined by testing the ability of the chemical in question to cause cytogenetic changes in the bone marrow of rodents. Traditional tests for carcinogenicity have ultimately relied on the use of laboratory animals, although this has obvious ethical implications.

The required time for a tumour to develop is known as the latent period. Some carcinogens induce tumours in the majority of experimental animals after a single dose within the latent period; however, the time lapse may be decreased by administering multiple doses of the chemical.

Diet and nutrients, obesity, alcohol and cancer

Diet, including some specific nutrients, obesity (*Chapter 10*) and alcohol are all known to be related to the incidence of different types of cancers.

Diet

The World Cancer Research Fund and the International Agency for Research on Cancer (IARC) have estimated that approximately 40% of cancer cases could be prevented by eating a healthy diet, maintaining an appropriate body weight (*Chapter 10*), participating in adequate physical activity and refraining from tobacco abuse.

Compelling evidence suggests a link between the overconsumption of foods of high glycaemic index (see *Margin Note 10.1*) and an increased incidence of cancer. The consumption of low fibre highly processed foods has a well-established association with the incidence of colorectal cancer (*Section 17.8*), although the link may be more complex than was initially thought. Indeed, an increased consumption of fibre-rich foods, such

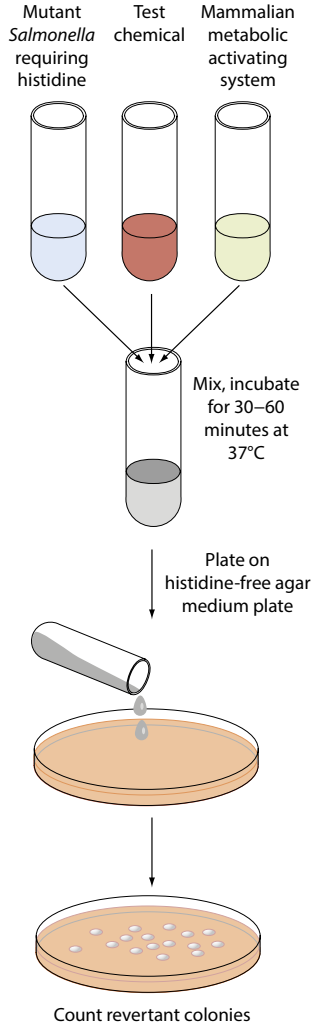

Mutant *Salmonella* requiring histidine
Test chemical
Mammalian metabolic activating system
Mix, incubate for 30–60 minutes at 37°C
Plate on histidine-free agar medium plate
Count revertant colonies

Figure 17.18 Schematic of the Ames test.

as fresh fruit and vegetables, has been correlated with a reduced incidence of several types of cancer, including those of the mouth, oesophagus, lung and stomach, in addition to those of the colon and rectum. Such associations have led to the recommendation by the UK Office for Health Improvement and Disparities and the NHS that individuals should consume at least five appropriately sized portions of a variety of fruit and vegetables daily. Cruciferous vegetables, such as cabbage, cauliflower, broccoli and Brussel sprouts, contain sulphoraphane (*Figure 17.19*), a chemical with anticancer properties. Increased consumption of such vegetables is inversely related to the incidence of prostate, breast, colon and bladder cancers.

Several nutritional interventions appear to offer some protection against developing cancer. These include decreasing the consumption of red and processed meats, the intake of which is linked to colorectal cancer. Increasing the consumption of ω-3 fats (see *Figure 10.7*), such as α-linolenic acid (ALA), eicosapentaenoic acid (EPA) and docosahexaenoic acid (DHA), is also recommended. Flax seeds, which are a good source of both ALA and dietary fibre, have been shown to reduce the development of carcinogen-induced tumours and to reduce the rates of metastasis in animal models.

Numerous studies using cancer cells and experimental animals have indicated that an appropriate vitamin D (*Chapter 10*) nutritional status offers protection against a variety of cancer types. It stimulates cellular differentiation and apoptosis (*Box 17.3*), but inhibits cell growth, reduces angiogenesis and decreases tumour metastasis. Vitamin D suppresses inflammation and modulates immunological functions (*Chapter 4*). However, inconsistent results were reported from several randomised trials on the association between vitamin D supplements and a reduced incidence of cancer and the associated mortality. This necessitates further research on the links between vitamin D concentrations or use of vitamin D supplements and the risk of developing or dying from cancer.

Extensive preclinical *in vitro* and *in vivo* studies over the past two decades have demonstrated that the diarylheptanoid curcumin (*Figure 17.20*), which is isolated from the plant turmeric (*Curcuma longa*), possesses anticancer properties. It interferes with multiple cell signalling pathways, induces cell cycle arrest, activates caspases (see *Box 17.3*), downregulates the production of antiapoptotic gene products and modulates oxidative stress. Currently, numerous clinical trials are investigating the therapeutic potential of curcumin-containing products to treat cancers, for example advanced pancreatic, breast and colorectal cancers. Some studies have suggested that antioxidants, such as those found in green tea, protect against cancers developing; however, other investigations have cast doubt on these observations. A high consumption of lycopene (*Figure 17.21*), a carotenoid found in tomatoes, has been reported in several studies to be associated with a decreased incidence of prostate cancer.

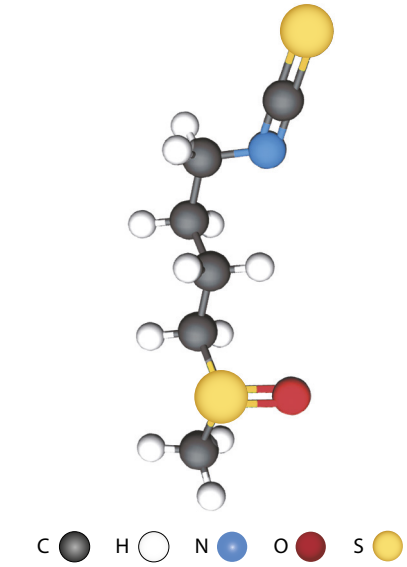


Figure 17.19 Computer drawn model of sulphoraphane.

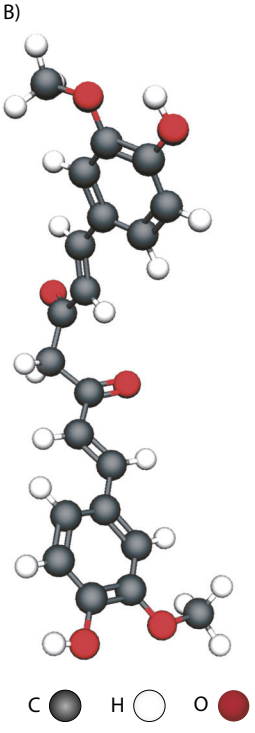


Figure 17.20 (A) Whole and ground rhizomes of turmeric (*Curcuma longa*) and (B) computer drawn model of curcumin.

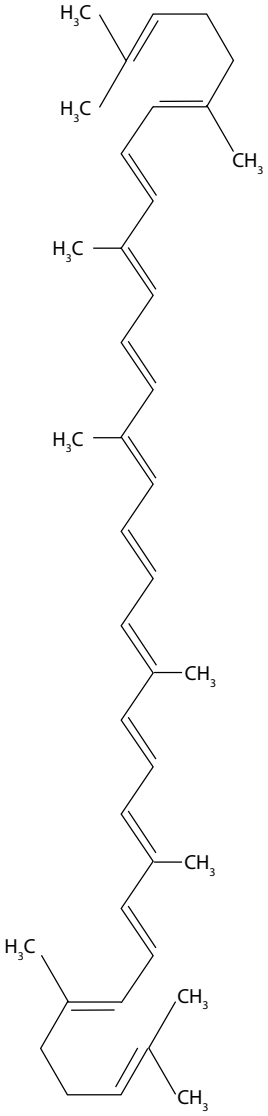


Figure 17.21 The structure of lycopene.

Obesity

Approximately 4–8% of all cancers are attributed to obesity (*Chapter 10*). Obesity has been strongly linked to the development of cancers of the breast, gallbladder, GIT, liver, kidneys, pancreas, and ovary and uterus. Studies indicated that overweight or obese women have a two to four times greater risk of developing endometrial cancer than those of healthy body weight.

Several possible mechanisms have been suggested to explain the link between obesity and the increased risk of cancer. Obese adipose tissue produces excessive quantities of proinflammatory cytokines, hormones, lipid metabolites and several adipokines, which may contribute to tumour initiation and progression (see *Chapters 4* and *7*). For example, excess fat accumulation causes an increased production of oestrogens, which can promote epithelial cell proliferation in the uterine endometrium, ovaries and mammary glands, and may, therefore, increase the risk of cancer developing in these tissues.

Obese individuals often have chronic inflammatory conditions that can induce oxidative stress and DNA damage, therefore increasing the risk of cancer. People with obesity often have higher circulating concentrations of insulin and insulin-like growth factor-1 (IGF-1; see *Figures 1.9* and *7.26*). Many types of cancer cells express increased levels of the insulin receptor. The binding of insulin to its receptor on cancer cells can activate those signalling pathways that promote cell growth and proliferation, potentially contributing to tumour growth.

Obesity alters the tumour microenvironment (*Box 17.5*), which can have implications for cancer progression. Adipose tissue in obese individuals undergoes significant changes, including extracellular matrix (ECM) remodelling, increased inflammation and fibrosis, which can create a supportive environment for tumour growth and progression.

Obesity not only is associated with an increased risk of cancer but may also increase the risk of cancer recurrence and mortality in cancer survivors. Hence, an early intervention to manage the weight of obese patients with early-stage cancer is necessary to improve the clinical outcome.

Alcohol (ethanol)

In 2020, 4.1% (6.1% and 2.0% in men and women, respectively) of all new cancer cases worldwide were attributed to the consumption of alcohol, strictly ethanol in all its commercial forms. The IARC has classified alcoholic drinks as a group 1 carcinogen to humans. Indeed, there is evidence of an alcoholic causality in breast, oral and pharyngeal, laryngeal, oesophageal, liver and colorectal cancers. Evidence is also accumulating that alcohol consumption is associated with an increased risk of melanoma, prostate and pancreatic cancers. The WHO has indicated that there is no safe limit on alcohol consumption and the risk of developing cancers.

Plausible mechanisms for alcohol-induced carcinogenesis may relate to its metabolism. The ethanol metabolite ethanal (formerly called acetaldehyde; *Chapter 12*) can cause DNA damage and block DNA synthesis and repair, while both ethanol and ethanal can disrupt DNA methylation. Ethanol can also induce inflammation and oxidative stress leading to lipid peroxidation and further DNA damage. Excessive ethanol consumption can cause folate and vitamin B deficiencies but increase oestrogen concentrations, which is significant in breast carcinogenesis.

Radiation and cancer

Ionising radiation can promote the transformation of cultured cells *in vitro* and the production of tumours *in vivo*. Supporting evidence for radiation-induced carcinogenesis comes from studies of Japanese people irradiated during the atomic bombings in Hiroshima (*Figure 17.22*) and Nagasaki, and from populations irradiated in nuclear accidents, such as that which occurred in Chernobyl in 1986. Long-term studies of Japanese atom bomb survivors showed an increase in the incidence of leukaemias in the first 5–10 years following exposure. The risk of solid tumours in these people was also increased significantly. An increased incidence of thyroid cancer has been reported in the decades following the Chernobyl accident in people who were children and adolescents at the time of the accident. Studies by Doll (*Box 17.6*) showed that infants previously subjected to X-irradiation *in utero* had an increased risk of developing leukaemias and solid tumours. A recent study of nuclear workers in France, the UK and the USA suggested a linear increase in the relative rate of cancers with increasing exposure to low-dose (0–100 mGy) ionising radiation.

The condition xeroderma pigmentosum also provides evidence for radiation-induced carcinogenesis. Exposure to UV radiation from sunlight induces damage to DNA, which

Figure 17.22 Hiroshima: the Peace Memorial Park. ⏎

is normally repaired through the nucleotide excision repair (NER) pathway. This pathway is responsible for recognising damaged DNA and repairing it by incision, ligation and biosynthetic steps. Eight *XP* genes encode proteins involved in different steps of the NER pathway. However, patients with XP are extremely sensitive to sunlight and are unable to mend the UV-damaged DNA because of mutations in these genes. People with XP have a 10 000-fold increased risk of developing skin cancer including basal cell carcinoma, squamous cell carcinoma and melanoma. They also have a 2000-fold increased risk of cancer of the eye. These symptoms appear early in life, typically before the age of 10 years.

Ionising radiation, X-rays, γ rays, α and β particles and neutrons cause DSBs and single-stranded breaks (SSBs) in DNA, and deletions, inversions and translocations in chromosomes (*Chapter 15*). Irradiation is also known to induce gene amplification and to increase chromosomal instability and these, in turn, increase the likelihood of mutations occurring. Ultimately, mutation events involving proto-oncogenes and/or tumour suppressor genes are the most likely causes of radiation-induced carcinogenesis. With some exceptions, cells are more susceptible to radiation-induced damage when dividing, and this fact has been utilised in the use of radiation to treat cancer (*Section 17.7*).

Viruses and cancer

Viruses were first implicated in the development of some types of cancer when F. Peyton Rous (1879–1970) showed in 1911 that a malignant tumour was caused by a 'filterable agent' (*Chapter 2*). This virus, which causes sarcomas in chickens, is now called the Rous sarcoma virus and has been used extensively in research into oncogenic, or cancer-inducing, viruses. Viruses are associated with 10% of all human cancers. In the 1960s, Burkitt lymphoma, a tumour of the jaws of children in certain regions of Africa, was found to be induced by a virus later identified as the Epstein–Barr virus (EBV; *Figure 17.23*) or human herpes virus 4 (see *Table 3.1*), which also causes glandular fever or infectious mononucleosis. This virus infects epithelial cells and B lymphocytes, causing their transformation, and is associated with several cancers, including nasopharyngeal carcinoma and Hodgkin lymphoma, in addition to Burkitt lymphoma.

Cancer-associated viruses belong to several families (International Committee on Taxonomy of Viruses [ICTV]), particularly prominent are retroviruses and papillomaviruses (*Table 17.7*).

Family of virus	Virus names and abbreviation(s)	Associated cancer(s)
Flaviviridae	Hepatitis C virus (HCV)	Hepatocellular cancer
Hepadnaviridae	Hepatitis B virus (HBV)	Hepatocellular cancer
Orthoherpesviridae	Epstein–Barr virus (EBV)/human gamma herpes virus 4 (HHV-4)	Burkitt and Hodgkin lymphomas; nasopharyngeal carcinoma
	Kaposi sarcoma-associated herpes virus (KSHV)/human gamma herpesvirus 8 (HHV-8)	Kaposi sarcoma; lymphoma
Papillomaviridae	Human papilloma viruses (HPV)	Cervical, vaginal, vulvar, penile, anal and oropharyngeal cancers
Polyomaviridae	Merkel cell polyomavirus (MCPyV)	Merkel cell carcinoma
Retroviridae	Human T cell leukaemia virus type 1 (HTLV1)/human T lymphotropic virus type 1	Adult T cell leukaemia/lymphoma

Table 17.7 Some viruses associated with human cancer. See also *Tables 3.1* and *3.3*.

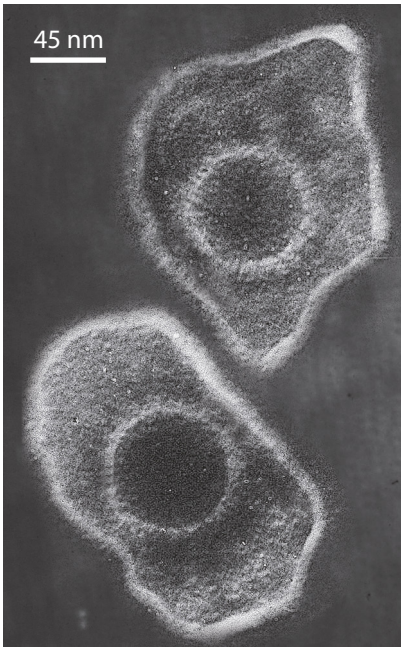


Figure 17.23 Electron micrograph of two Epstein–Barr virus/human herpes virus 4 virions. The virus capsids are loosely surrounded by their envelopes. Adapted from (2005) Virus proteins prevent cell suicide long enough to establish latent infection. *PLoS Biol* 3(12):e430. doi: 10.1371/journal.pbio.0030430.

Retroviruses are RNA-containing viruses that have reverse transcriptase activity (*Chapters 2* and *3*). When they infect a cell, the reverse transcriptase transcribes their RNA to give a complementary DNA, which may become incorporated into the host genome. Not all retroviruses cause cancer, but those that do are called 'transforming retroviruses'. The human immunodeficiency virus (HIV) is a retrovirus that is also associated with cancer, although in this case the association is indirect. People with HIV

who develop full-blown acquired immunodeficiency syndrome (AIDS; *Chapter 3*) have an increased incidence of tumours, such as lymphomas and Kaposi sarcoma, which are associated with EBV and Kaposi sarcoma-associated herpes virus (KSHV). It is likely that the immunosuppression caused by HIV allows these latent viruses to escape immunological control.

The most widely recognised association between viruses and cancer occurs with certain strains of human papilloma viruses (HPVs) that are linked to the development of cervical cancer (*Figure 17.24*). More than 200 subtypes of HPV have been found. The genomes of these viruses all have similar structures. They contain at least six 'early' genes (E1, E2, E4, E5, E6 and E7) and two 'late' genes, L1 and L2 (*Figure 17.25*). The early genes encode proteins responsible for HPV genome replication, cellular transformation and proliferation, and immune modulation. The late genes encode L1 and L2 capsid proteins that have roles in viral assembly and the transmission of the virus. The nonprotein coding part of the genome contains the upstream regulatory region (URR) necessary for the transcription and replication of viral DNA. The HPV E6 and E7 genes are viral oncogenes: the proteins they encode inactivate the tumour suppressor proteins p53 and RB, respectively (*Section 17.4*). Thus, cell proliferation is induced, apoptosis is prevented and DNA damage accumulates. Additionally, E6 activates human telomerase reverse transcriptase (*Box 17.1*) in keratinocytes, allowing telomere lengths to be maintained and the cells to become essentially immortal.

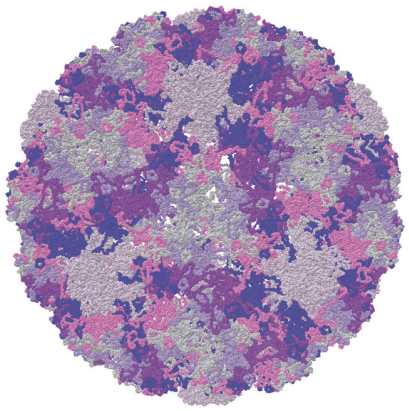


Figure 17.24 Molecular model of the capsid of the human papilloma virus (HPV), strain 16. PDB file 3JBA. ⏎

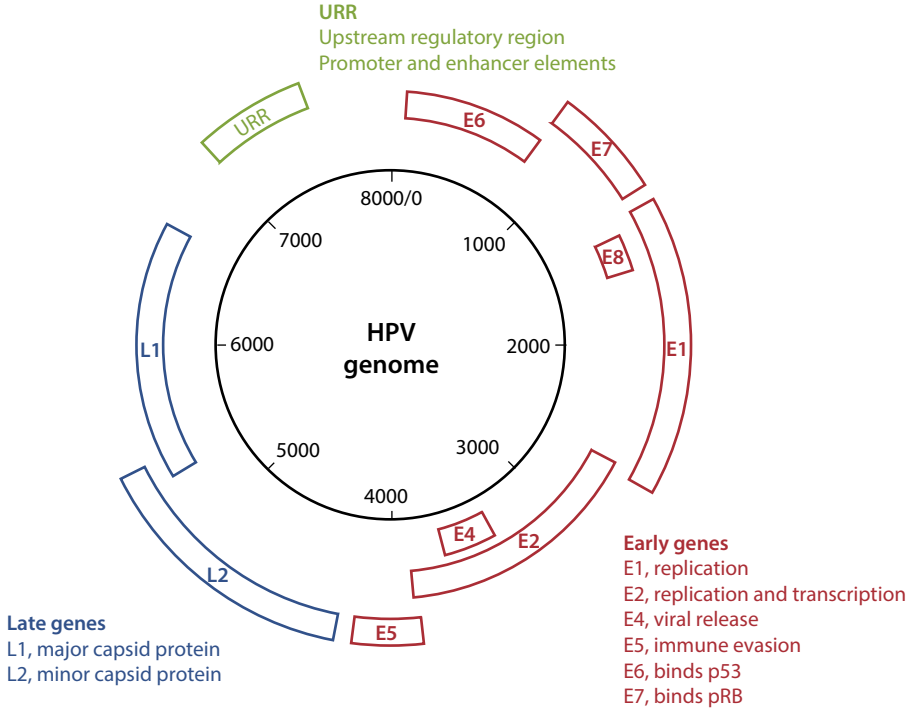


Figure 17.25 Simplified genome of a human papilloma virus (HPV). ⏎

Based on their oncogenic potential, HPVs are classified into low-risk and high-risk categories. Some low-risk HPVs, such as HPV1, 2 and 4, invade epithelial cells of the skin, while others, for example HPV6 and 11, infect the genital tract causing benign warts, but with a low risk of cancer. Unfortunately others, such as HPV 16, 18, 31, 33, 35, 39, 45, 51, 52, 56, 58 and 59, belong to the high-risk category and are associated with the development of cervical cancer. These 12 HPVs are currently classified as carcinogenic by the IARC. DNA from at least one of these high-risk types is detected in approximately 99% of human cervical cancers. Human HPV 16 and 18 infections are responsible for nearly 70% of cervical cancers.

Cervical cancers develop from precursor 'lesions' that are graded low to high according to how much disruption of epithelial cell differentiation has occurred (*Figure 17.26*). In cervical carcinomas, viral DNA becomes integrated into the host genome. However, many benign lesions also contain these strains of HPV, so the presence of the virus alone is insufficient to cause cancer. It seems that long-term infections with high-risk HPVs predispose the individual to cervical cancer and that other factors, for example the carcinogens in cigarette smoke, immunosuppressive drugs and some of the older oral contraceptive pills, can increase the risk of precancerous cervical cells progressing to full malignancy.

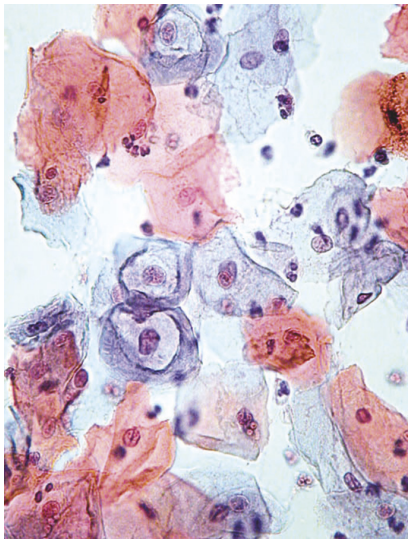


Figure 17.26 Photomicrograph of a cervical smear showing red-stained squamous cells from the superficial layer of the cervix and darkly stained cells from the layer of the cervical wall immediately below that of the squamous cells. Note that these cells are abnormal in having comparatively large nuclei and show that the patient is at risk of developing cancer of the cervix. See also *Figure 1.22*. Courtesy of H. Glencross, Manchester Cytology Centre, Manchester Royal Infirmary, UK. ⏎

In addition to cervical cancer, HPV infections are also associated with other cancers (*Table 17.7*). In 2019, the WHO estimated that HPVs caused 620 000 and 70 000 cases of cancer in women and men, respectively. It has been estimated that worldwide 31% of men over the age of 15 were infected with at least one type of genital HPV in 2023. Screening can detect cervical precancers that can be treated before they become cancerous, but vaccination (*Chapters 2* and *3*) is the best way to prevent HPV infections, cervical cancer and other HPV-related cancers. In the UK, the HPV vaccine has been part of the NHS routine vaccination schedule since 2008. It is recommended for children aged 12–13 years old and for people at higher risk of HPV. The HPV vaccine used in the NHS vaccination programme is Gardasil 9. Since 2023, it has been given in a single dose, and protects against nine types of HPV strains: 6, 11, 16, 18, 31, 33, 45, 52 and 58.

Margin Note 17.6 Cervical screening programmes

The UK NHS cervical screening programme is available to women aged 25–64 years to take a smear test to check the health of their cervix every three to five years. The aim of the screening is to identify any HPV infection, find precancerous cervical cell changes and, if necessary, allow early treatment to prevent cervical cancer from developing. A small sample of cells is taken from the cervix and checked for high-risk types of HPVs. If these types of HPV are not found, there is no need for any further tests. If HPVs are found (HPV positive), but abnormal cell changes have not developed, then the person will be invited for another screening the following year. If the sample is positive for HPV and abnormal cells are also present (see *Figure 17.26*), a colposcopy, a detailed examination of the cervix and associated anatomical areas, and a biopsy test are recommended as the next investigative stages.

17.6 GENERAL DIAGNOSIS OF CANCER

The symptoms of cancer depend on the type of tumour, its location and the extent of its development. Symptoms can develop when the cancer is relatively small, for example when even a small tumour of the brain exerts pressure that results in clinical lesions. Other tumours, such as one developing in an ovary, may not produce symptoms until they are relatively large.

A number of general signs may indicate the presence of an undiagnosed tumour, for example an unintentional weight loss of more than 5% of the body weight in a year. Weight loss that occurs because of cancer is due to the release of cytokines from cells of the immune system. In end-stage cancer, this weight loss is known as cachexia (*Box 17.7*). Fever, which is also induced by cytokines, is often found in patients with advanced cancer, although this may also be an early warning of certain types of cancer. An example of this is Hodgkin lymphoma, which is characterised by fever, often during sleep, and is accompanied by drenching night sweats. Fatigue may also be a general symptom, particularly where the cancer causes a loss of blood with concomitant anaemia, as may occur, for instance, in stomach cancers. Other signs and symptoms include the presence of a lump in a breast or testicle, or unusual bleeding or discharges.

BOX 17.7 CACHEXIA

Cachexia (derived from the Greek *kakos* meaning bad and *hexis*, condition) is characterised by a severe loss of weight, mainly from the loss of adipose and skeletal muscle tissues, in someone who is not actively trying to lose weight. It is marked by fatigue, weakness, ill health and decreased appetite. It is usually a sign of various underlying disorders and is associated with chronic diseases such as cancer and chronic infectious diseases such as tuberculosis (*Chapter 3*), or occurs secondarily because of an inability to ingest or use nutrients, for example from an obstruction in the gastrointestinal tract (GIT) or clinical malabsorption (*Chapter 11*). Surgery or treatment-related disorders, such as the nausea and vomiting associated with chemotherapy or radiation therapy, may also be involved. Approximately half of all cancer patients lose weight, although the extent varies with the type of cancer. At diagnosis, 80% of patients with upper gastrointestinal cancers and 60% with lung cancer are already suffering cachexia, although patients with haematological malignancies and breast cancer usually avoid such a substantial loss of weight. Most other solid tumours are associated with a higher frequency of cachexia.

Cachexia is more common in children and elderly patients and becomes more pronounced as the disease progresses. Its prevalence increases from 50% to over 80% before death. In addition to compromising a patient's quality of life, it is correlated with poor outcomes. Indeed, cachexia is the main cause of death in over 20% of patients, often due to degeneration of the respiratory muscles or cardiac arrhythmias and chronic heart failure following atrophy of cardiac muscle (*Chapter 14*).

In cancer patients, both tumour and immune system cells secrete numerous cytokines (*Chapter 4*): cachectic factors that result in metabolic dysfunction(s) causing tumour-associated cachexia. These molecular signals include interleukins 1 and 6 (IL-1, -6), interferon-γ (IFN-γ) and tumour necrosis factor α (TNF-α), also called cachectin, which stimulate a pronounced inflammation and the production of acute phase proteins by the liver, and the breakdown of proteins and lipids (*Figure 17.27*). The release of proteolysis-inducing factor also promotes protein catabolism, with the release of amino

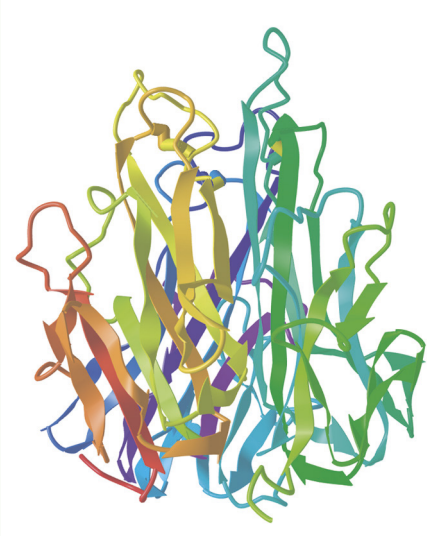


Figure 17.27 Molecular model of cachectin (tumour necrosis factor α [TNF-α]). PDB file 1TNF.

acids. Lipid breakdown is stimulated by the secretion of a lipid-mobilising factor, called zinc α2-glycoprotein (ZAG), from tumour cells (*Figure 17.28*), which promotes lipolysis in white adipose tissue (WAT). The released amino acids and free fatty acids are recycled by the liver to supply glucose, the major metabolic fuel used by cancer cells. The liver also recycles the lactate released by the tumour. Further, the conversion of WAT to brown adipose tissue (BAT) can occur. Unlike WAT, BAT does not simply store energy but catabolises lipid reserves to generate heat. This combination of metabolic effects means cancer patients have high basal metabolic rates, and the degradation of lean muscle and adipose tissues produces the cancer-associated cachexia. Biochemical signalling between patient and cancer cells also leads to neuroendocrine malfunctioning. Some signals have leptin-like effects on the hypothalamus, depressing the appetite and contributing to a severe cancer-linked anorexia (*Chapter 10*); patients also tend to become insulin resistant (*Chapter 7*), which further exacerbates the cachexia.

Cachexia can be diagnosed in a patient by evaluating his or her nutritional status, usually with a combination of clinical assessment and anthropometric tests as described in *Chapter 10*. Body weight, with reference to the normal adult weight, is the usual measure of nutritional status.

Naturally, the best way to treat cancer cachexia would be to cure the cancer. Unfortunately, in adult patients with advanced solid tumours, this is often not possible, particularly when the condition has been diagnosed at an advanced stage. Treatment is therefore generally aimed at improving the quality of life and, for many patients, this means improving appetite and food intake, encouraging physical exercise where possible and trying to inhibit muscle and fat wasting. Unfortunately, treatment is limited. Hypercaloric feeding, either enteral or parenteral, does not generally increase lean mass. Glucocorticoids are widely used as palliatives because they inhibit the biosynthesis and/or release of proinflammatory cytokines, such as TNF-α and IL-1, which decrease food intake directly, or through other anorexigenic mediators, such as leptin (*Chapter 10*). Corticotrophin-releasing factor (CRF) and serotonin have some limited effect in improving appetite and food intake. Corticosteroids, for example dexamethasone, hydrocortisone and prednisolone (*Chapter 6*), have significant antinausea effects and improve asthenia (weakness), reduce inflammation and help in pain control. However, prolonged treatment can lead to weakness, delirium, osteoporosis and immunosuppression, all of which are significant problems anyway in advanced cancer patients. Megestrol acetate (MA), medroxyprogesterone acetate (MPA) and ω-3 fatty acids, such as eicosapentaenoic acid (see *Figure 10.7*), all improve appetite, energy intake and nutritional status.

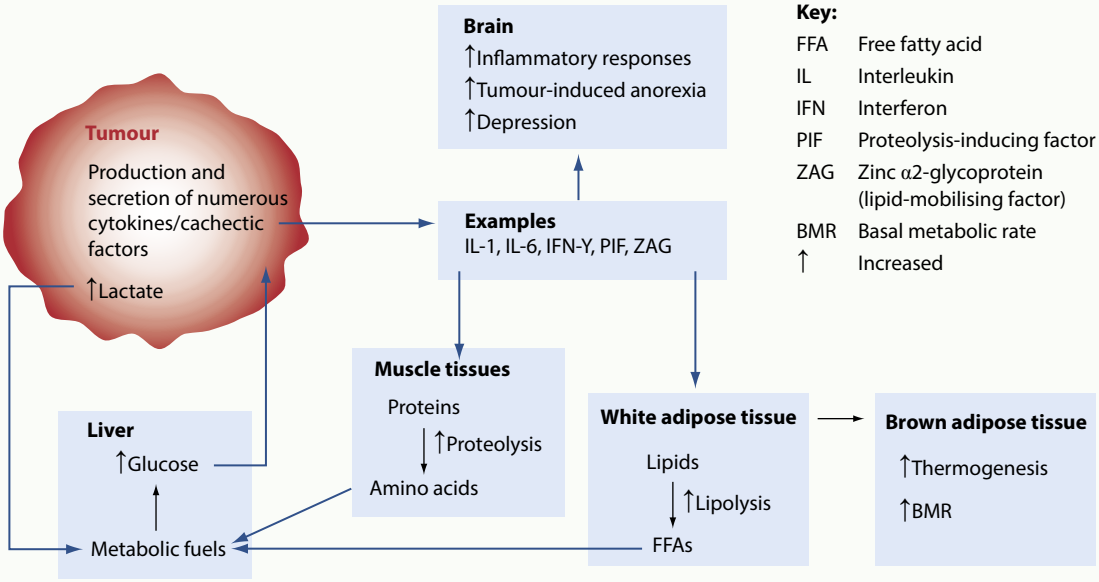


Figure 17.28 Some of the biochemical/physiological activities involved in cancer-associated cachexia.

Clinical tests for cancers

Clinical tests for cancer are used to screen 'at-risk' populations, to detect cancer in a patient presenting with indicative symptoms, to monitor the success of treatment and to detect recurrence in a patient who has been in remission. These tests may involve the detection and/or quantification of tumour-associated molecules (known as tumour markers/biomarkers), the detection and localisation of tumours within the body, and the cytological or histological examination of biopsy material to detect precancerous conditions and/or determine the nature of any tumour present.

Detection and estimation of tumour markers

Tumour markers are genes, proteins or other substances produced by cancer cells or, indeed, healthy body cells whose presence or increased concentrations are associated

with the development of specific types of malignant tumours or certain benign tumours. They can be tested for in samples of blood, serum, urine, stools, body fluids like pleural fluid, tumours, or other tissues such as lymph nodes and bone marrow. Analysing tumour markers can provide a range of information that is necessary for cancer care, such as helping diagnose the type or subtype of a malignancy, staging the extent of tumour spread (*Sections 17.7* and *17.8*), determining the most advantageous therapy, monitoring therapeutic responses, predicting prognosis and examining for recurrence, and screening at-risk populations.

Table 17.8 lists examples of tumour markers used clinically. Some are associated with only a single type of cancer, whereas others are associated with multiple types. Some are used for diagnosis, others mainly to determine and monitor treatment.

Tumour marker	Examples of associated cancer
BCR-ABL fusion gene (Philadelphia chromosome; see *Figures 17.12* and *17.13*)	Acute and chronic lymphocytic leukaemias; acute and chronic myeloblastic leukaemias
BCL2 gene rearrangement	Lymphomas, leukaemias
CA19-9[a]	Bile duct, gallbladder, gastric and pancreatic cancers
CA125[a]	Ovarian cancer
Calcitonin	Medullary thyroid carcinoma
Carcinoembryonic antigen (CEA)	Breast, colorectal and lung cancers
Human chorionic gonadotropin (hCG)	Choriocarcinoma and germ cell tumours
EGFR mutation	Colorectal and non-small cell lung cancers
α-Fetoprotein (AFP)	Hepatocellular and ovarian cancers; teratoma of testis
PIK3CA mutation	Breast, colorectal and non-small cell lung cancers
Prostate-specific antigen (PSA)	Prostate cancer

[a] Cancer antigen.

Table 17.8 Examples of tumour markers and their associated cancers. ⏎

Some conventional tumour markers are 'tumour-associated antigens', for example cancer antigens 19-9 and 125 (CA19-9, CA125), carcinoembryonic antigen (CEA) and prostate-specific antigen (PSA; *Section 17.8*). The concentrations of CA125 and PSA are increased in the bloods of female and male patients with cancers of the ovaries and prostate gland, respectively. Carcinoembryonic antigen is a glycoprotein that is overproduced in most carcinomas of the breast, colon and lung. These markers can be detected and quantified using monoclonal antibodies (MAbs) specific for the antigen in question (see *Box 4.5*). The antibody may be used in an enzyme-linked immunosorbent assay (ELISA) or in radioimmunoassay (RIA) as described in *Box 4.2*.

Unfortunately, increases in tumour-associated antigens are not exclusive to cancers, so the tests can only give an indication of cancer. Increased concentrations of an antigen should therefore always be followed by further tests, such as diagnostic imaging.

Detection and localisation of tumours within the body

Patients who present with symptoms indicative of cancer may be examined using diagnostic imaging procedures to localise the potential tumour or to determine the extent of metastasis. Diagnostic imaging may involve the use of X-rays, computed tomography (CT), magnetic resonance imaging (MRI), positron emission tomography (PET) and ultrasound, all of which are undertaken by clinical radiologists.

Computed tomography scanning uses X-rays to take sequential pictures of the body from different directions. In practice, the patient may be required to drink a contrast solution, or this may be administered intravenously, to enhance the tissue contrast. The patient then lies still on a table that passes through the X-ray machine, which rotates around the patient taking pictures of thin 'slices' of tissue. The images are combined computationally to produce three-dimensional images or computed tomograms.

Magnetic resonance imaging and PET, outlined in *Chapter 18*, may also be used to visualise abnormal growths within the body and to determine the extent of tumour growth. Ultrasound (*Chapter 14*) may be used to locate tumours within the abdomen, including tumours of the liver and ovaries.

In the UK, routine screening for breast cancer by mammography is offered to women over the age of 50 years. A tumour will show up as a 'shadow' on a breast mammogram. Patients who have lung cancer may have X-rays taken of their bones to determine whether the tumour has spread to them. However, to determine whether the abnormal growth is a cancerous or benign lesion, it is essential to examine biopsies of the growth.

Biopsies and histology

Biopsy material may be obtained from a variety of sources in one of several ways. Samples from solid tumours may be obtained by endoscopy or during surgical procedures involving local or general anaesthesia. An *endoscope* is a relatively long thin flexible tube with a camera and light at the end. Depending on the type of tumour, the endoscope is inserted into a body cavity and allows internal tissues such as the GIT and the lungs to be viewed. Endoscopes also enable samples of the suspect tissue to be removed without the need for surgery. In *bronchoscopy*, an endoscope is used to obtain samples of lung tissue in suspected cases of lung cancer (*Figure 17.29*; *Section 17.8*).

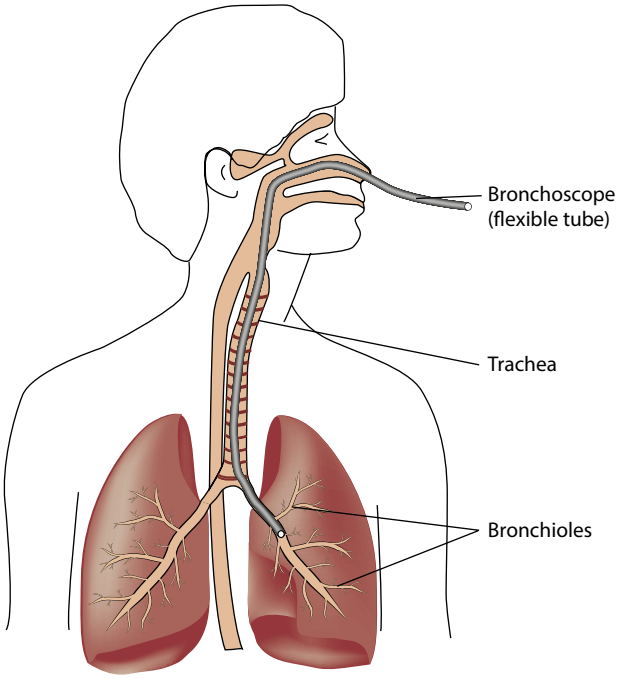


Figure 17.29 Schematic showing the use of an endoscope in bronchoscopy.

Needle biopsies collect small amounts of tissue from a variety of solid tissues. Samples of blood and bone marrow may be examined to detect leukaemias (*Section 17.8*). Also, some solid tumours may cause an accumulation of fluid, for example ascites fluid in the peritoneal cavity (*Chapter 10*) or fluid in a pleural effusion in the thorax (*Chapter 14*), which can be examined microscopically to determine if they contain suspended tumour cells and so diagnose a neoplasm. Finally, some precancerous lesions may be detected by removing smears of tissue and examining them histologically, for example in the preparation of cervical smears (see *Figures 1.22* and *17.26*) in order to detect precancerous lesions as outlined below.

Solid tissue obtained for histological examination must be processed to obtain thin sections for microscopic examination. The material is first preserved by fixation in, for example, formalin, dehydrated and then embedded in paraffin wax to support the tissue. This allows sections of 2–7 μm thick to be cut using a microtome. The sections are mounted on slides, dewaxed with xylene, rehydrated and then stained in the manner appropriate to the investigation. The most common stains for tissue sections are haematoxylin and eosin, which stain the nucleus blue/purple depending on the formulation and cytoplasm pink, respectively. Thus, these stains provide an excellent contrast within the cells (*Figure 17.30*). Smears of blood and bone marrow are most frequently

Figure 17.30 Section of skin stained with haematoxylin and eosin. Note the germinative layer of rapidly dividing cells, which is damaged by many anticancer drugs, as indicated by the arrowhead.

stained with May–Grünwald–Giemsa, which stains nuclei blue, while the cytoplasmic contents are stained differentially depending on the cell type. This type of staining therefore allows different types of leukocytes to be identified. The stain commonly used to detect precancerous cells in cervical smears is the Papanicolaou stain, or Pap stain, as it is more frequently known. This formulation contains five different stains: haematoxylin, which stains the nucleus; Orange G (OG-6) and EA, which contain light green; eosin Y; and Bismarck brown Y. Orange G and EA are counterstains. The Pap stain is often used to stain buccal and sputum smears, as well as those obtained from the cervix.

The preparation of material from solid tissues for histological examination is an extended process. Sometimes a more rapid examination may be required, for example when the result of a biopsy is needed during an operation to determine the extent of surgery required. In such cases, the biopsy is rapidly frozen to –176°C in liquid nitrogen. This process hardens the material and allows it to be cut into 5–10 μm thick sections using a freezing microtome or cryostat. Once the sections have been prepared, they can be stained with haematoxylin and eosin without the long procedures required with paraffin sections, often within minutes of removal from the body. Examination will then, hopefully, ensure the appropriate surgical procedure.

Cancer cells can be specifically stained if they bear a tumour-specific marker, such as CEA (see above). Cryostat sections are often used for this process because fixation, wax embedding and clearing can destroy some antigenic sites on the tumour. The sections are stained by immunohistochemistry. In this process, the sections are incubated with a MAb to a tumour-associated antigen. This specific binding is visualised by incubation with an enzyme-labelled anti-immunoglobulin, followed by a further incubation with the appropriate substrate. For enzyme immunohistochemistry, a colourless substrate is chosen that produces a coloured insoluble product (see *Box 4.2*). Antigenic sites that were destroyed during the preparation of wax embedded sections can be 'retrieved' using enzymatic/proteolytic or heat-induced antigen retrieval methods. In the enzymatic method, proteases such as proteinase K, trypsin and pepsin are used. The heat-induced method uses heat (from microwaving, pressure cooker, steamer, water bath or autoclave) and a selection of buffers. The optimal antigen retrieval technique is dependent on the antigen, the tissue, the fixation method and/or the primary antibody used. Some antigens require a combination of enzymic digestion and heating.

Molecular diagnosis

The histological diagnosis of potentially cancerous cells is increasingly supported by molecular diagnostic techniques. Molecular methods are used frequently to test genomic markers, such as tumour gene mutations, patterns of tumour gene expression and nongenetic changes in tumour DNA, across a range of tumours. The determination of the genetic profile of a primary tumour can inform therapy, including personalised medicine (see *Margin Note 17.7*), and guide prognosis. One molecular way of investigating a tumour requires biopsy material to be extracted using a fine needle and its DNA amplified using the polymerase chain reaction (PCR; *Chapter 3*). The DNA is then analysed for mutations known to be implicated in cancer by one of several methods. One widely used diagnostic method is fluorescence *in situ* hybridisation (FISH; *Figure 17.31*).

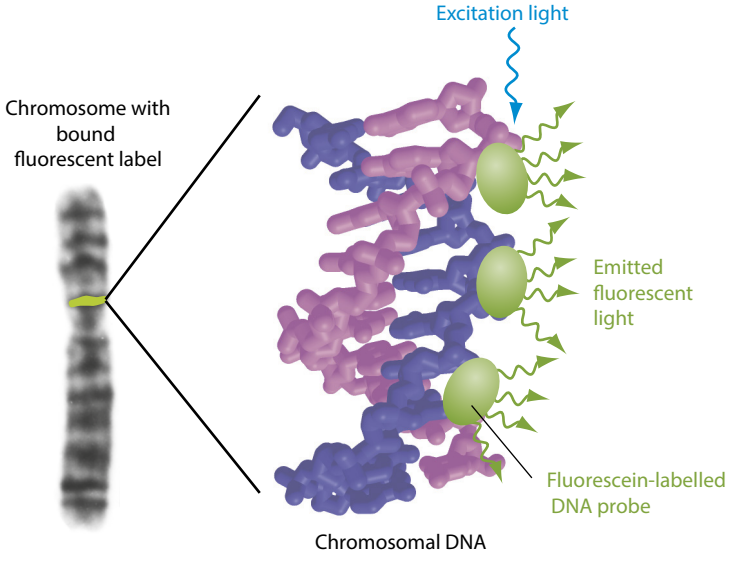


Figure 17.31 Schematic outlining the binding of a fluorescently labelled DNA probe to target specific chromosomal DNA sequences (fluorescence *in situ* hybridisation [FISH]). See also *Figures 1.16*, *5.10* and *17.56*.

Thin sections of the tumour are treated to denature the DNA, which is then hybridised *in situ* with fluorochrome-labelled probes for relevant mutated, cancer-associated genes. The slides are then examined using a fluorescence microscope. Mutations are indicated by fluorescent spots in the nuclei (see *Figures 1.16*, *5.10* and *17.56*).

Tissue biopsies are the most widely used tool for detecting cancers, but sometimes tumour tissue can be difficult to obtain, especially in metastatic diseases like late-stage lung cancer.

A liquid biopsy is a blood test that detects circulating tumour cells (CTCs) and/or circulating tumour DNA (ctDNA) fragments. These methods can detect cancers at an early stage of their development, allowing therapy to also start beneficially earlier. Circulating tumour cells are ones that have detached from the primary tumour and entered the circulatory or lymphatic systems; they have the potential to form metastatic tumours. Thus, the presence of CTCs can be an early indicator of cancer metastasis. Cell-free tumour DNA is released from cells by apoptosis, necrosis and autophagy, or by active release mechanisms. The biggest hurdle in CTC detection is the extremely low number of cells in blood. However, the concentration of approximately 1 CTC/mL blood in samples can be increased by exploiting their physical properties, such as size and density, to enrich the samples. The development of NGS technologies (*Box 17.2*) means sequencing ctDNA is also highly sensitive. Hence, the analysis of ctDNA fragments can provide information about cancer cell genomes (somatic mutations) and epigenetic changes. Compared with traditional cancer diagnosis using tissue biopsy, liquid biopsy is less invasive and a more comprehensive way to evaluate tumour heterogeneity. It can also help plan treatment, assess how well treatment is working or determine if a cancer has returned to a patient who had been in remission. Indeed, liquid biopsy methods have revolutionised cancer diagnosis and treatment.

The first liquid biopsy test approved by the United States Food and Drug Administration (FDA) in 2016 identified metastatic non-small cell lung cancer (NSCLC; *Section 17.8*) in patients who would therefore be eligible for treatment with erlotinib. The method analyses samples of plasma to detect a deletion in exon 19 or a substitution in exon 21 mutations in the epidermal growth factor receptor (*EGFR*) gene. Additional liquid biopsy tests have now been approved by the FDA. The number of genes that can be analysed in liquid biopsy samples is remarkable. One test can analyse over 300 genes, including the tumour markers *BRCA1* and *BRCA2*, *ATM* alterations, *PIK3CA* variations, *ALK* rearrangement and the *EGFR* exon 19 deletion and exon 21 substitution, to help diagnose breast, non-small cell lung, ovarian and prostate cancer patients who may benefit from targeted therapy (*Margin Note 17.7*). Another test detects mutations in the *BRAF*, *EGFR*, *IDH1*, *HER2*, *ROS1* and *RET* genes in patients with bile duct, non-small cell lung and thyroid cancers, and cholangiocarcinoma.

17.7 GENERAL TREATMENT OF CANCERS

The first treatments for primary tumours without metastasis were surgical excision of the neoplasm, followed by chemotherapy or radiation therapy (radiotherapy), or attempts to stimulate an immune response to fight the cancer. Chemotherapy involves the use of drugs, while radiotherapy uses ionising radiation to kill cancer cells. However, other types of cancer treatment now available include hormone therapy, immunotherapy, and targeted and gene therapies. Patients with cancer may have only one type of treatment or a combination of them. Often, palliative care is offered to patients regardless of the stages of their cancers or their prognosis. The aims of palliative care are to improve the quality of life by alleviating symptoms, easing pain and discomfort, and addressing the patient's physical, emotional, mental and social needs as they transition to the end of life.

Staging is performed following a diagnosis of cancer. It involves determining the extent of development of the cancer. Staging is essential, as it determines the treatment the patient receives. A simple staging system classifies the tumour into one of five stages.

- Stage 0: the tumour has not spread to nearby tissues; it is called carcinoma *in situ* (CIS).
- Stages I–III: the larger the tumour, the bigger the number assigned and the more extensive its spread into nearby tissues and/or the lymph nodes.
- Stage IV: the cancer has metastasised to at least one other organ/system.

The TNM system devised by the American Joint Committee on Cancer (AJCC) and the Union for International Cancer Control (UICC) is the most widely used cancer staging system. The 'T' defines the size and extent of the primary tumour growth in or from its original location, 'N' describes the tumour's spread to local lymph nodes and 'M' describes whether the tumour has metastasised. The TNM staging varies between different types of tumours. Most hospitals and medical centres use TNM systems, and

Margin Note 17.7 ⏎ Personalised medicine

Personalised medicine, also called precision medicine, is an approach in which clinical care is tailored to meet the individual patient needs, that is, to prevent, diagnose or treat any disease. It is based on information regarding the individual patient's genes, proteins and other biochemicals. It has, therefore, only become possible with the development of rapid and inexpensive next-generation sequencing methods (see *Box 17.2*). In cancer therapy, personalised medicine also considers the medical history, lifestyle and family medical history, which are integrated with the clinical data to form a diagnosis and develop the patient's specific, that is, personalised, therapy. The effectiveness of the treatment is continuously assessed and used in prognosis.

examples are shown in *Tables 17.15* and *17.17*. Some cancers, such as those that affect the central nervous system (CNS) and leukaemias (*Section 17.8*), cannot use the TNM system for classification. Instead, the tumour grade system can be used for brain tumour or other solid tumour classification, which describes how normal or abnormal cancer cells appear when examined microscopically. A tumour grade (*Table 17.9*) reflects how closely the tumour cells resemble their cells of origin.

Grade	Description
X	Grade cannot be assessed (undetermined grade)
1	Well-differentiated cells (low grade tumour)
2	Moderately differentiated cells (intermediate grade tumour)
3	Poorly differentiated cells (high grade tumour)
4	Undifferentiated abnormal cells (high grade tumour)

Table 17.9 Grading of tumours. ⏎

The type of cancer, its stage and grade, and the patient's general condition are crucial factors in determining the appropriate treatment. After evaluating the potential benefits and risks of each treatment option, a cancer treatment team, which may include oncologists, surgeons, radiologists, pathologists, pharmacists, nurses and other specialists, will develop a comprehensive treatment plan that addresses all aspects of patient care. The goals of treatment can vary from patient to patient.

Chemotherapy

A substantial number of drugs are used to treat different cancers, given singly or in combinations (*Chapter 3*), or with other types of therapy. Most of the drugs used in the treatment of cancers are cytotoxic and are most active on dividing cells, that is, they are cell cycle dependent. Drugs that kill cells that are not dividing are noncycle dependent.

Drugs that target dividing cells do so by a variety of mechanisms. For example, several interfere with DNA biosynthesis. These include alkylating agents, antimetabolites such as folic acid antagonists, purine and pyrimidine analogues, and inhibitors of topoisomerases (*Chapter 3*). Other drugs inhibit cell division by disrupting the polymerisation or depolymerisation of microtubules, thus interfering with the separation of chromosomes during mitosis (*Chapter 15*). Unfortunately, drugs that work by preventing cell division do not discriminate between cancerous and healthy cells and display considerable cytotoxicity, particularly towards progenitor cells in the bone marrow and epithelial cells of the skin and GIT. Hence, chemotherapy is usually associated with anaemia, nausea and damage to the actively dividing cells of the hair follicles, leading to considerable, but usually reversible, hair loss. One role of hospital pharmacists is to advise on treatments that minimise the discomfort caused by this therapy. In the treatment of some cancers, aggressive chemotherapy may destroy bone marrow to the extent that patients require haematopoietic stem cell transplantation (HSCT; *Chapters 5 and 6*).

The route for administrating chemotherapeutic agents depends on the drug. Some are given orally, while others are injected into muscle tissue, the fat tissue under the skin, the peritoneal cavity or directly into an artery that supplies the cancer. Many are administered intravenously by infusion. Others are given intrathecally, that is, by injection into the innermost membrane surrounding the CNS, usually by lumbar puncture. Administration of chemotherapy in hospitals requires the hospital pharmacist to work alongside the clinician to ensure the most appropriate dose is provided in the most suitable manner. Some of the drugs used to treat cancer are discussed below, although this list is by no means exhaustive. It is worth noting that a natural selection process often takes place within tumours treated with chemotherapy, in that some of the tumour cells may develop resistance to the drug. In cases of drug resistance (*Margin Note 17.8*), it is necessary to change the chemotherapeutic agent(s). New drugs are continually being developed and approved for cancer treatment.

Folic acid antagonists

Dihydrofolate reductase (DHFR) is active in the synthesis of tetrahydrofolate, which is required for the biosynthesis of purines and pyrimidines, which are precursors of nucleotides, RNA and DNA. Folic acid antagonists, such as methotrexate, structurally resemble the substrate for DHFR (*Figure 17.32*). They bind to DHFR and inhibit its activity.

Margin Note 17.8 ⏎
Drug resistance in cancer

One major concern associated with cancer treatment is drug resistance. Although many cancers are initially susceptible to chemotherapy or targeted therapy, over time they can develop resistance that significantly compromises the effectiveness of treatment. Drug resistance in cancer cells can be caused by decreased drug uptake, enhanced drug efflux, altered drug metabolism, improved DNA damage repair, genetic and epigenetic changes of drug targets, and cancer cell plasticity. These mechanisms act either individually or in combination, leading to the development of single or multidrug-resistant cancer cells. Note how several of these mechanisms have also evolved in the resistance to antibiotics shown by numerous bacteria (see *Box 3.2*).

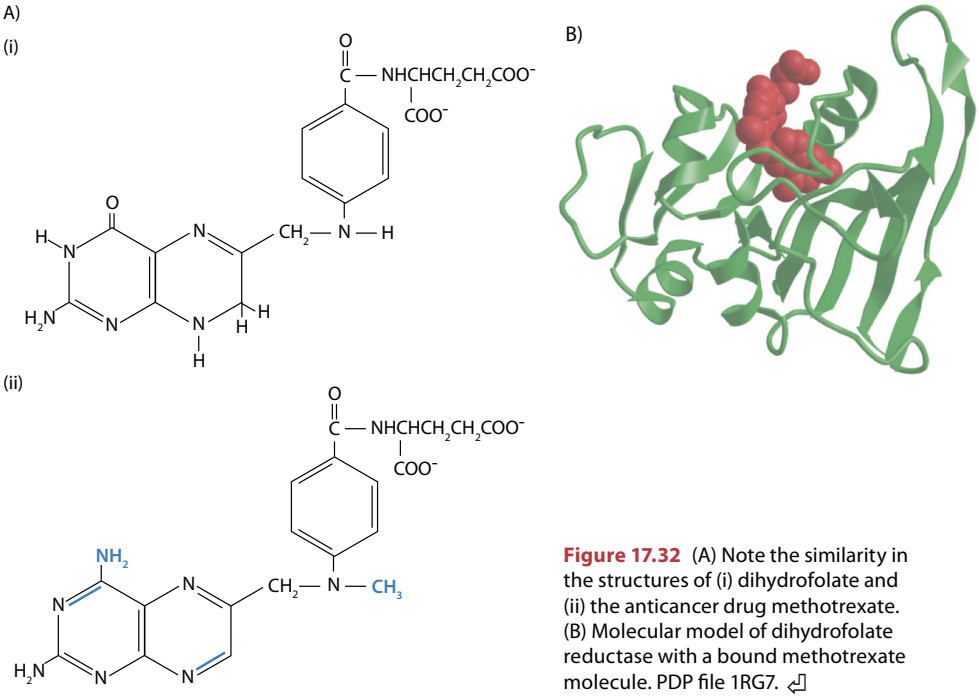


Figure 17.32 (A) Note the similarity in the structures of (i) dihydrofolate and (ii) the anticancer drug methotrexate. (B) Molecular model of dihydrofolate reductase with a bound methotrexate molecule. PDP file 1RG7. ⏎

Margin Note 17.9
Repurposing drugs to treat cancers

Drug repurposing, also called drug repositioning or drug reprofiling, is the process of identifying new therapeutic uses for existing drugs that are already approved for other indications. Usually, *de novo* drug discovery and development takes 10–17 years, with a low overall probability of success, while repurposing a drug is estimated to take only six to seven years. Drug repurposing potentially reduces the expense and also the risks associated with development, as safety, dosing and pharmacokinetic data are already available for the repurposed drug. Repurposing existing drugs for cancer treatment is an area in cancer research that has had several notable successes. For example, thalidomide, which was initially developed as a sedative and antinausea medication, has been repurposed for the treatment of multiple myeloma. The reverse has also occurred: drugs developed for cancer treatment have been repurposed for alleviating organ rejection in transplant surgery (*Chapter 6*).

Alkylating agents

Alkylating agents are electrophiles that can covalently bind to DNA bases. Some alkylating agents are monofunctional and react with only one strand of DNA. Others are bifunctional and react with both strands, covalently linking them together. Alkylating agents work in all phases of the cell cycle by preventing the separation of the strands that is necessary for the biosynthesis of new DNA, which is the major cause of the toxicity associated with these drugs. Examples of alkylating agents, such as nitrogen mustards and the nitrosoureas used in treating cancers, are shown in *Table 17.10*.

Class	Drug	Examples of cancers treated
Nitrogen mustards	Bendamustine	Chronic lymphocytic leukaemia, B cell non-Hodgkin lymphoma
	Chlorambucil	Chronic lymphocytic leukaemia, Hodgkin lymphoma
	Cyclophosphamide	Leukaemia, lymphomas
		Adenocarcinomas of breast and ovary, neuroblastoma, retinoblastoma
	Ifosfamide	Testicular cancer
	Melphalan	Multiple myeloma, ovarian cancer
Nitrosoureas	*Bis*-chloroethyl nitrosourea (BCNU)	Brain tumours, Hodgkin lymphoma, non-Hodgkin lymphoma, multiple myeloma
	Cyclohexyl-chloroethyl nitrosourea (CCNU)	Brain tumours, Hodgkin lymphoma
	Streptozocin	Pancreatic cancer

Table 17.10 Example of alkylating drugs used to treat cancers. ⏎

Purine and pyrimidine analogues

Purine and pyrimidine analogues are drugs that structurally resemble the bases found in DNA and/or RNA. Their use interferes with the biosynthesis of DNA in the S phase of the cell cycle, although the sites at which they exert their effects depend on the drug itself (*Figure 17.33*).

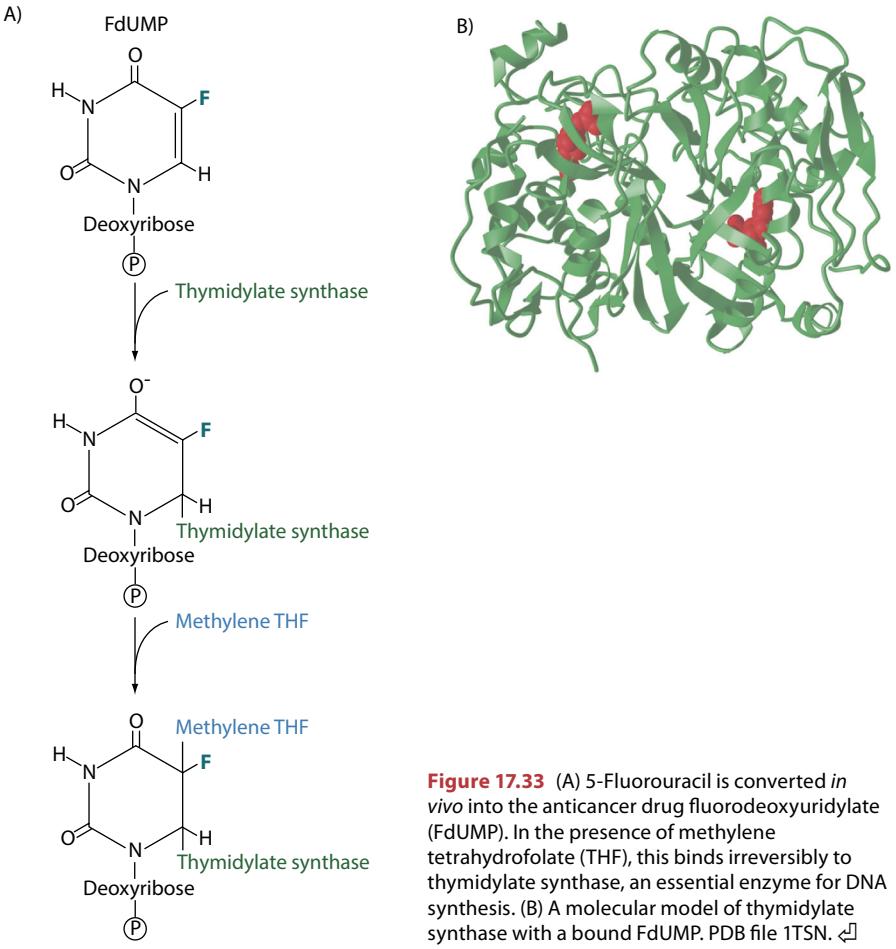


Figure 17.33 (A) 5-Fluorouracil is converted *in vivo* into the anticancer drug fluorodeoxyuridylate (FdUMP). In the presence of methylene tetrahydrofolate (THF), this binds irreversibly to thymidylate synthase, an essential enzyme for DNA synthesis. (B) A molecular model of thymidylate synthase with a bound FdUMP. PDB file 1TSN. ⏎

Examples of pyrimidine and purine analogues and some of their uses in cancer treatment are listed in *Table 17.11*.

Class	Drug	Examples of cancer treated
Pyrimidine analogues	5-Fluorouracil (5-FU)	Breast, cervical, head and neck, and gastrointestinal cancers
	Cytidine arabinoside (Ara-C)	Acute leukaemia
	Cladribine	Hairy cell and chronic lymphocytic leukaemias
	Gemcitabine	Breast, non-small cell lung, ovarian and pancreatic cancers
Purine analogues	6-Mercaptopurine (6-MP)	Acute lymphoblastic and acute myeloid leukaemias
	6-Thioguanine (6-TG)	Acute lymphoblastic and acute myeloid leukaemias

Table 17.11 Examples of pyrimidine and purine structural analogues used in cancer treatment. ⏎

Antimicrotubular agents

Antimicrotubular agents interfere with the formation of the mitotic spindle, which is essential for chromosome separation during mitosis and meiosis. The formation of a spindle requires the polymerisation of tubulin subunits, while its dissolution requires their depolymerisation. Any drug that interferes with either of these processes will arrest mitosis in metaphase and inhibit cell division. The first drug shown to prevent polymerisation was colchicine, which was originally obtained from the autumn crocus

(*Colchicum autumnale*). It is also used as an anti-inflammatory agent to ameliorate acute gout. Research into its potential therapeutic use for cancer continues. The vinca alkaloids vincristine, vinblastine and vinorelbine, which were all derived originally from the periwinkle (*Vinca rosea*), work in a similar manner to colchicine (*Figure 17.34*). Vinblastine is used in combination with other drugs to treat breast and testicular cancers, choriocarcinoma, lymphoma and Kaposi sarcoma. Vincristine is used in the chemotherapy of leukaemias, lymphoma, neuroblastoma, rhabdomyosarcoma and Wilms tumour. Vinorelbine is used in lung and breast cancer therapy. Cabazitaxel, docetaxel (Taxotere) and paclitaxel, otherwise known as taxol, are derived from the bark and needles of the Pacific yew tree (*Taxus brevifolia*; see *Figure 15.7*). These drugs also target microtubules but by preventing their depolymerisation. Both paclitaxel and docetaxel (*Figure 17.35*) are used in the treatment of breast, lung and ovarian cancers. Cabazitaxel is used to treat castration-resistant metastatic prostate cancer (*Section 17.8*).

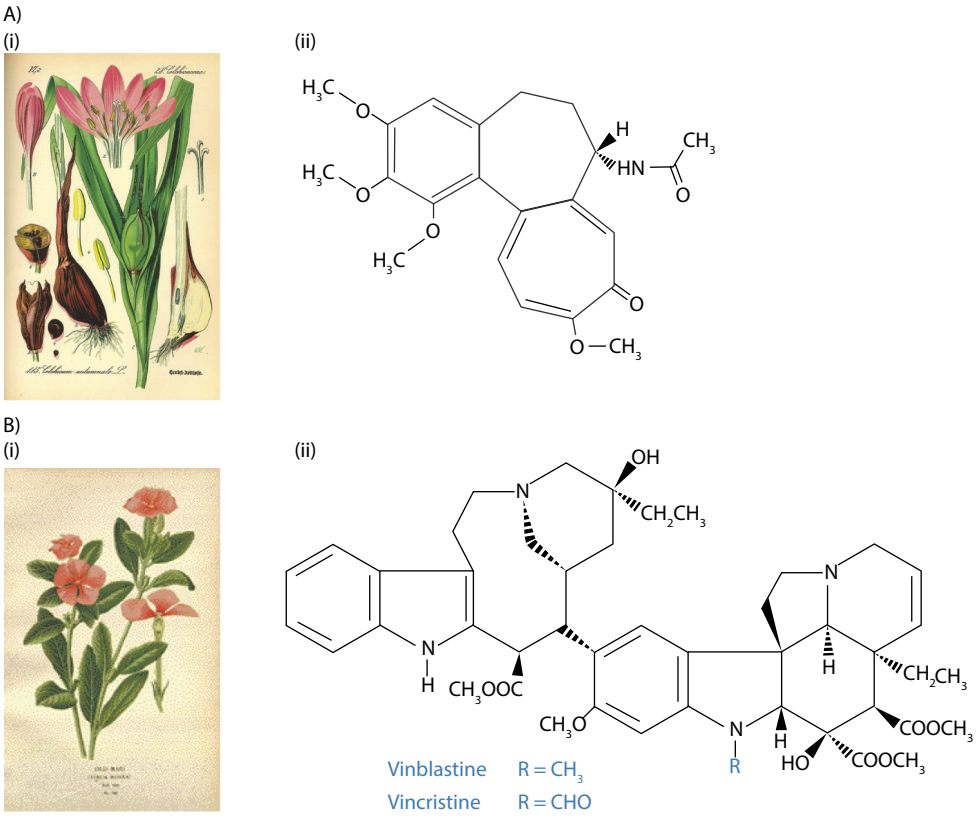


Figure 17.34 (A) (i) The autumn crocus (*Colchicum autumnale*), the original source of (ii) colchicine. (B) (i) The periwinkle (*Vinca rosea*) and (ii) the vinca alkaloids vinblastine and vincristine.

Topoisomerase inhibitors

Topoisomerase inhibitors, such as irinotecan (topoisomerase I inhibitor), etoposide, teniposide and anthracyclines (topoisomerase II inhibitors), inhibit the biosynthesis of DNA by forming a complex with the DNA and the topoisomerase at S or G2 phases of the cell cycle. The complex induces breaks in the DNA and prevents their repair by inhibiting topoisomerase activity. Accumulated breaks in the DNA prevent cells from entering the mitotic phase, and eventually lead to cell death. Etoposide and teniposide are derivatives of podophyllotoxin, which is derived from the mandrake plant (*Madragora*). Etoposide is used to treat bladder, small cell lung and testicular cancers, and lymphomas; teniposide to treat childhood leukaemia; and irinotecan in the chemotherapy of colorectal cancer (*Section 17.8*).

Daunorubicin and doxorubicin, both isolated from cultures of *Streptomyces peucetius* var. *caesius*, are anthracycline antibiotics. Doxorubicin is recognised as one of the most efficient chemotherapy medications approved by the FDA for the treatment of a variety of cancers, such as those affecting the blood (leukaemias), breast, lung, ovaries and thyroid, and also sarcomas. Daunorubicin is especially useful in the treatment of leukaemias. Dactinomycin is also an antibiotic derived from *Streptomyces parvullus*. It binds to DNA and inhibits its transcription and therefore protein synthesis. Dactinomycin is used clinically to treat Ewing sarcoma, rhabdomyosarcoma, Wilms tumour and testicular cancer.

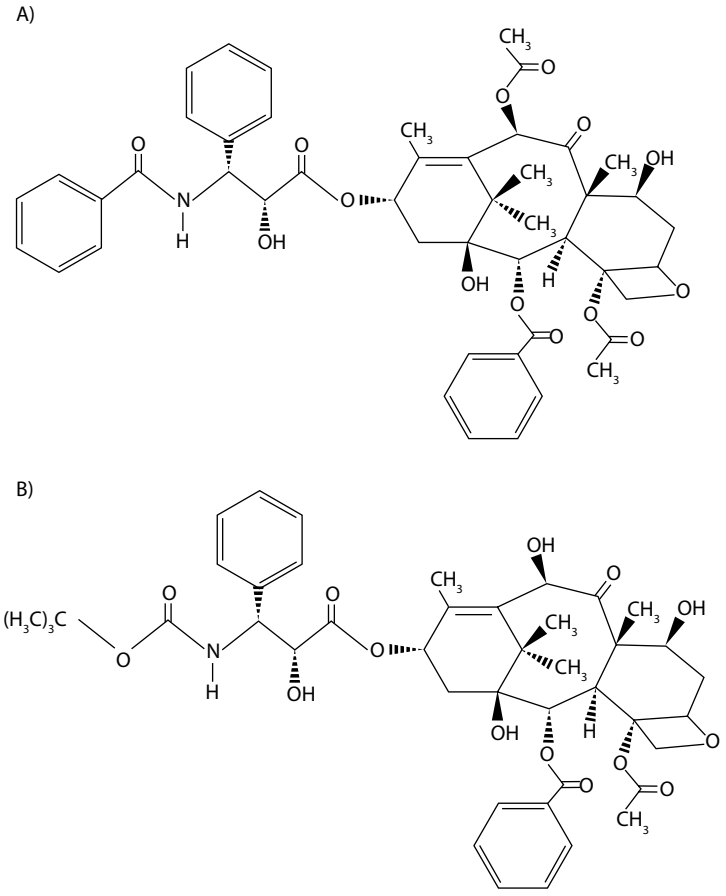


Figure 17.35 Structures of (A) paclitaxel (taxol) and (B) docetaxel (Taxotere).

Platinum-based drugs

Cisplatin (*cis*-diamminedichloroplatinum) was discovered in experiments that showed the growth of bacteria was inhibited when platinum electrodes were used to deliver an electric current. It is now used in combination therapy to treat bladder, lung, ovarian and testicular cancers, and a number of sarcomas. Other platinum-based drugs include carboplatin and oxaliplatin. All function by binding to DNA, cross-linking the strands and distorting the double helix (*Figure 17.36*). This distortion facilitates the binding of other anticancer drugs to the DNA that improves their therapeutic effects.

Radiation therapy

Radiation therapy or radiotherapy involves the use of high-energy electromagnetic waves to kill cancer cells, as most dividing cells are susceptible to radiation. It is used to treat about 50% of patients with cancer worldwide. Common side effects of radiation therapy are fatigue, nausea and some external burns to the skin; thus, the dose of radiation must be carefully calculated to give the optimum antitumour dose but cause minimal side effects.

Radiation treatment is delivered in one of two ways:

1. external beam therapy
2. internal radiation therapy

In external beam therapy, as the name implies, the area is irradiated with electrons, photons (X-rays or γ rays) or protons from an external source. Electrons cannot travel far through tissues; therefore, their use is limited to tumours on the skin or near the surface of the body. Most radiation therapies use photon beams that can irradiate the tumour within the body. External beam therapy uses images from CT, MRI and PET scans to plan the session(s) of treatment. This includes adjustments the position(s) of the radiation source, and dose and duration of the session. Patients will usually be treated by radiation therapy once a day but have several sessions. Treatment regimens vary depending on the size and location of the tumour and the purpose of the treatment, that is, whether

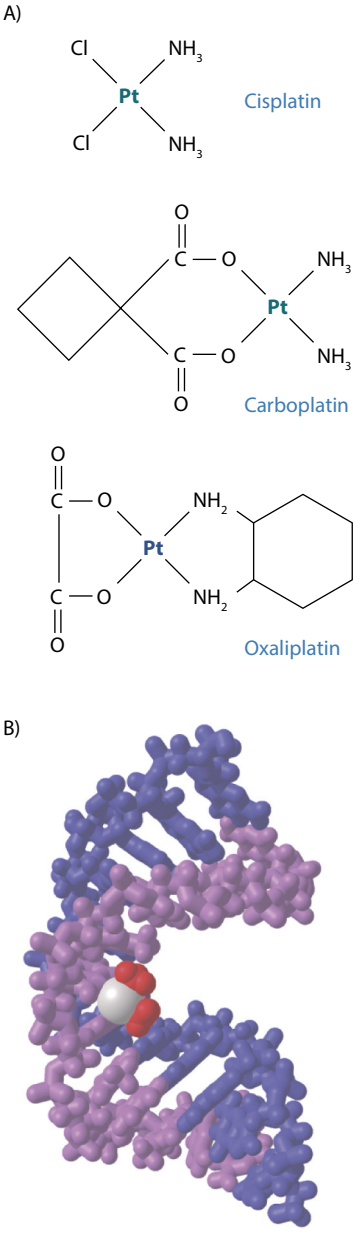


Figure 17.36 (A) Structures of the platinum-based anticancer drugs cisplatin, carboplatin and oxaliplatin. (B) A molecular model of cisplatin bound to DNA. Note how the double-helical structure has been bent and distorted by the binding of the drug. PDB file 1A84.

it is hoped to cure the patient, to shrink the tumour prior to surgery or chemotherapy, or to palliate an incurable tumour.

Internal therapy involves placing a radioisotope (*Table 17.12*) near or inside the tumour for a short period of time, a process known as brachytherapy. For example, intracavitary radiotherapy involves the insertion of ^{192}Ir in an applicator into a body cavity. This type of treatment is used for cancers of the vagina, cervix and uterus. Alternatively, thin radioactive wires may be inserted directly into the tumour, for example in treating prostate cancer. Internal therapy may also involve giving the patient a radioactive liquid either orally or intravenously. For example, a drink containing radioactive ^{131}I may be given to treat thyroid cancer. Since the thyroid preferentially takes up iodine (*Chapter 7*), the ^{131}I becomes concentrated at the site where it is required. Intravenous radioactive liquids are also used to treat metastatic bone cancer.

Isotope	Mean photon energy (*KeV*)	Half-life (days)	Examples of cancers treated
Caesium-131 (^{131}Cs)	30	9.7	Brain, head, neck, lung, prostate and other localised tumours
Iodine-125 (^{125}I)	28	59.4	Abdominal, brain, head, neck, pelvis, prostate, spine and thoracic cancers
Iodine-131 (^{131}I)	192	8.1	Thyroid cancer
Iridium-192 (^{192}Ir)	380	73.8	Breast, cervical, head, lung and neck cancers
Palladium-103 (^{103}Pd)	21	17	Prostate cancer
Rhenium-188 (^{188}Re)	784	0.71	Nonmelanoma skin cancer
Yttrium-90 (^{90}Y)	940	2.7	Liver cancer and non-Hodgkin lymphoma

Table 17.12 Examples of radioactive isotopes used in internal radiation therapy to treat cancers.

Proton beam therapy

Proton beam therapy or proton radiotherapy is a type of cancer treatment that, as the name implies, uses protons (hydrogen ions, H^+) as the therapeutic radiation. A synchrotron or cyclotron accelerates the H^+, increasing their energy to sufficient levels that they can penetrate to the desired depth in the body and destroy cancer cells. Proton beam therapy has fewer side effects than other types of radiation therapies, such as using X-rays, because it can be targeted more directly at the tumour and does less damage to surrounding tissues. It can be used alone or combined with surgery, chemotherapy or immunotherapy. It is often used for treating children with certain cancers of the brain and spinal cord, such as malignant meningioma, retinoblastomas and orbital rhabdomyosarcomas, because it lessens the chance of harming healthy, growing tissues. Given that proton therapy uses expensive technology, the equipment is only available at a limited number of cancer centres (*Figure 17.37*).

Figure 17.37 The Proton Beam Therapy Centre at The Christie Hospital, Manchester. Courtesy of The Christie NHS Foundation Trust, Manchester, UK.

Hormone therapy

Hormone therapy, also known as endocrine therapy, is a type of cancer treatment that involves removing, blocking or adding hormones to the body to slow down or stop the growth of cancer. Hormone therapy is often used along with other cancer treatments, for example shrinking a tumour before surgery or radiation therapy, preventing cancer recurrence after initial treatment or treating metastatic cancers. Hormone therapy is primarily used to treat hormone receptor-positive breast, prostate and endometrial (uterine) cancers that depend on sex hormones to grow.

Examples of hormone therapies used in cancer treatment are summarised in *Table 17.13*. Their mechanisms of action are diverse. For example, aromatase inhibitors (such as the steroid exemestane) and nonsteroidal agents (for example anastrozole; see *Figure 17.42*) inhibit cytochrome P-450 aromatase (also called CYP19A1; oestrogen synthase; see *Figure 17.43*) activity, which reduces the concentration of circulating oestrogen in postmenopausal women with oestrogen-dependent (receptor-positive) breast cancer (*Section 17.8*). Tamoxifen (see *Figure 17.41*) and fulvestrant inhibit the binding of oestradiol to oestrogen receptors (ERs) and/or induce a conformational change in the

Margin Note 17.10 Huggins and hormone therapy

Around 1940, Charles Huggins (1901–1997) discovered that the growth of prostate cancer depends on hormones. He showed that testosterone stimulated and oestrogen decreased the activity of prostate cancer. His discovery led to the development of hormone therapy as a standard treatment for prostate cancer. Huggins also established phosphatase as a tumour marker to measure the disease response. He was awarded the Nobel Prize in Physiology or Medicine in 1966.

Types of hormone therapy	Example of medication	Cancer treated
Aromatase inhibitor	Anastrozole Exemestane Letrozole	[a]ER+ and/or PR+ breast cancer in postmenopausal women; ER+ and/or PR+ metastatic endometrial cancer
Oestrogen receptor modulator	Tamoxifen	[a]ER+ and/or PR+ breast cancer in both women and men
Oestrogen receptor antagonist	Fulvestrant Toremifene	[a]ER+ metastatic breast cancer in postmenopausal women
Luteinising hormone-releasing hormone (LHRH) agonists	Goserelin Leuprolide Triptorelin	[a]ER+ breast cancer in premenopausal women; advanced endometrial cancer; prostate cancer
Luteinising hormone-releasing hormone (LHRH) antagonist	Degarelix Relugolix	Advanced prostate cancer
Antiandrogen	Apalutamide Enzalutamide Darolutamide	Castration-resistant prostate cancer
CYP17[a] inhibitor	Abiraterone ketoconazole	Advanced prostate cancer with metastasis
Progestin	Medroxyprogesterone acetate Megestrol acetate	Endometrial cancer

[a] CYP17, 17α-hydroxylase, whose activity is essential for the biosynthesis of androgens.
ER+, oestrogen receptor positive; PR+, progesterone receptor positive.

Table 17.13 Examples of hormone therapies used in treating cancers.

receptor. This prevents the hormone–receptor complex binding to oestrogen-response elements in DNA (see, for example, *Figure 7.14*), and prevents the hormone-stimulated expression of oestrogen-dependent genes and therefore cell proliferation. In addition, tamoxifen upregulates the production of transforming growth factor B (TGFb). This factor inhibits the growth of tumour cells, and downregulates the expression of insulin-like growth factor 1 (IGF-1), a factor that stimulates breast cancer cell growth. Goserelin is a synthetic analogue of luteinising hormone-releasing hormone (LHRH). It binds to and activates pituitary gonadotropin-releasing hormone (GnRH) receptors. Thus, prolonged administration of goserelin inhibits the secretion of pituitary gonadotropin, thereby decreasing the concentrations of circulating testosterone (in males) and oestradiol (in females), which can result in the regression of sex hormone-sensitive prostate and breast tumours, respectively.

Immunotherapy

Immunotherapy in cancer therapy refers to the manipulation of immune system (*Chapter 4*) activities to improve its responses against a tumour. The use of immunotherapy has a long history and began before the advent of cytotoxic drugs. Examples of such relatively poor immunotherapy include the use of Coley's toxin. The toxin contained a mixture of killed *Streptococcus pyogenes* and *Serratia marcescens* that nonspecifically stimulated the immune system (*Chapter 4*). The BCG vaccine, which contains killed mycobacteria, was used in the 1960s and 1970s to treat malignant melanoma, and has since been used to treat non-muscle invasive bladder cancer. Mycobacteria are potent stimulators of the immune response and increase the production of several cytokines, including interferon γ (IFN-γ) and tumour necrosis factor α (TNF-α). Contemporary treatments use recombinant cytokines to enhance immune responses. For example, interferon α (IFN-α) has been used successfully in the treatment of multiple myeloma, chronic myelogenous leukaemia, hairy cell leukaemia (*Margin Note 17.17*), follicular non-Hodgkin lymphoma, Kaposi sarcoma and melanoma. Interleukin 2 (IL-2) enhances the proliferation and cytotoxicity of natural killer cells (NKCs); it has been used to treat melanoma and renal cancer (see *Figure 4.17*).

Cancer vaccines

Sipuleucel-T and talimogene laherparepvec (T-VEC) are two therapeutic cancer vaccines in current use. The former is applied to combat metastatic hormone refractory prostate cancer, while T-VEC is used to treat metastatic melanomas that cannot be removed surgically. Sipuleucel-T is an autologous cellular immunological agent, so it is 'tailor made' for each patient. Its production involves collecting antigen-presenting cells (APCs) from the patient (*Chapter 4*). These are then incubated with recombinant fusion protein antigen, which contains prostatic acid phosphatase (PAP) and granulocyte-macrophage colony-stimulating factor (GM-CSF). This process activates the APCs. The activated, antigen-loaded APCs are infused into the patient to attack the cancerous prostate cells and, hopefully, cure the patient. Talimogene laherparepvec is made using a genetically modified type I herpes simplex virus that expresses GM-CSF. It enhances the number of antigen loading major histocompatibility complex (MHC) class I molecules and so increases the tumour-antigen presentation capacity of dendritic cells. Thus, when T-VEC is injected directly into melanomas in the skin or lymph nodes, it lyses the tumour cells, releasing tumour-derived antigens that together with the virally derived GM-CSF promote an antitumour immune response.

Therapeutic mRNA-based cancer vaccines (*Chapter 3*) have yet to be approved for standard treatment; however, results from clinical trials are promising. Therapeutic mRNA cancer vaccines are more likely to succeed in combination with other immunotherapeutic treatments such as immune checkpoint inhibitors and adoptive cell therapy.

Immune checkpoint inhibitors

Immune checkpoints are immune cell surface receptors that are involved in the control of activation or inhibition of the immune response. They are a normal part of the immune system, crucial in maintaining self-tolerance, avoiding autoimmunity and preventing excess immune responses to minimise collateral damage to healthy tissues. The discovery of immune checkpoint proteins, such as the following, was a significant breakthrough in the field of cancer immunotherapy:

- programmed cell death protein-1
- programmed death-ligand 1
- cytotoxic T lymphocyte antigen 4
- lymphocyte activation gene-3

The programmed cell death protein-1 (PD-1; also known as CD279) is a transmembrane glycoprotein of M_r 55 000, which is expressed on activated T and B cells, NKCs and APCs. Its ligand, programmed death-ligand 1 (PD-L1; also known as CD274 or B7 homologue 1 [B7-H1]), is also a transmembrane glycoprotein (M_r 33 000) that is usually expressed by activated T and B cells, APCs and some epithelial cells, particularly in inflammatory conditions. When PD-1 recognises and binds to PD-L1 on an activated T or B cell, then T cell proliferation is inhibited, cytokine production suppressed and T cell exhaustion induced. The cytotoxic T lymphocyte antigen 4 (CTLA-4, CD152) is a glycoprotein of M_r 25 000 that is mainly expressed on activated T cells and regulatory T (T_{reg}) cells. The transmembrane protein lymphocyte activation gene-3 (LAG-3 or CD223; M_r 57 000) is found on the surface of activated T cells and NKCs. All three transmembrane proteins, PD-1, CTLA-4 and LAG-3, are so-called co-inhibitory receptors and negatively regulate T cell-mediated immune responses. Cancer cells can exploit these inhibitory molecules, for instance by overexpressing PD-L1 to induce tumour tolerance and T cell exhaustion, and therefore prevent the immune system from destroying them.

Immune checkpoint inhibitors (ICIs) can block the actions of checkpoint proteins and so reactivate the immune response against cancer cells. Examples of ICIs approved for use by the FDA to treat a variety of cancers include inhibitors of:

- PD-1 (cemiplimab, nivolumab and pembrolizumab)
- PD-L1 (atezolizumab, avelumab and durvalumab)
- CTLA-4 (ipilimumab and tremelimumab)
- LAG-3 (relatlimab)

In 2024, the National Institute for Health and Care Excellence (NICE) recommended that nivolumab and relatlimab be used for untreated unresectable or metastatic melanoma in patients aged 12 years or older.

Chimeric antigen receptor T cell therapy

Chimeric antigen receptor T cell (CAR-T) therapy uses T cells that have been isolated from the blood of a patient (*Figure 17.38*). The T cells are then genetically modified by

Margin Note 17.11
Immunotherapy innovators: Allison and Honjo

The concept of checkpoint therapy in cancer treatment is attributed to the work of James Allison (1948–) and Tasuku Honjo (1942–). They were jointly awarded the Nobel Prize in 2018 for their discovery of a new cancer treatment through the inhibition of negative immune regulation. Allison demonstrated that blocking cytotoxic T lymphocyte antigen 4 (CTLA-4), the negative immune regulator, was able to stop tumour formation in mice. Honjo discovered the programmed cell death protein-1 (PD-1) and showed that it had similar functions to CTLA-4, that is, it was also a T cell brake. Together, their discoveries revolutionised cancer treatment by harnessing the body's immune system to fight cancer cells.

inserting a gene that encodes a CAR, which can recognise and bind to a specific tumour-associated antigen (TAA) on the cancer cells of the patient. The genetically modified T cells are then grown to an appropriate density in tissue culture before being infused into the patient. Here, the CAR-expressing T cells can target and eliminate the TAA-expressing tumour cells.

Commercial CAR-T cell therapies used to treat multiple myeloma, B cell acute lymphoblastic leukaemia and large B cell lymphoma have been approved by the European Medicines Agency (EMA) and FDA since 2017. Their potential as treatment for other types of cancer is being studied. Acute toxicities, such as cytokine release syndrome (see *Margin Note 4.3*), are one of the challenges associated with the use of CAR-T cell treatments. However, the application of advanced gene-editing technologies, such as CRISPR gene editing (see below), means CAR-T cells can be engineered more efficiently to become safer and more effective as therapeutic agents.

Targeted therapy

Targeted cancer therapy employs drugs designed to specifically recognise unique molecular characteristics of cancer cells, allowing them to be distinguished from healthy body cells and subjected to therapy. Targeted therapies hold the promise of more focused personalised cancer treatments (*Margin Note 17.7*). It minimises side effects and enhances patient outcomes. Tumour markers (*Table 17.8*) that control cancer cell growth, division, survival and metastasis are major targets for this type of treatment. Drugs used in targeted therapies include small-molecule inhibitors and MAbs that can directly inhibit the expression of mutant genes or the actions of overactive/overabundant proteins. Some MAbs are engineered to carry toxins, together known as antibody–drug conjugates (ADCs), specifically to cancer cells; this approach combines the targeting specificity of a MAb with the cytotoxic effects of the chemotherapeutic drug. However, as with all cancer therapies, tumours can become resistant to drugs used. Thus, it may work best when used with more than one type of targeted therapy or when combined with other types of cancer treatment.

Examples of targets, available drugs and cancers are shown in *Table 17.14*.

Target	Genetic alterations	Drugs	Cancers treated
ATM	Mutation	Olaparib	Breast cancer
BCR-ABL	Fusion	Dasatinib, imatinib, nilotinib, ponatinib	Acute lymphoblastic leukaemias; chronic myelogenous leukaemia
HER2	Amplification, mutation	Lapatinib, neratinib, tucatinib, trastuzumab, pertuzumab	Breast cancer
K-RAS	Mutation	Sotorasib	Non-small cell lung cancer
PIK3CA	Mutation	Alpelisib	Breast cancer
PML-RARA	Fusion	Retinoic acid, arsenic trioxide	Myeloid leukaemia
RET	Mutation, fusion	Pralsetinib, Selpercatinib	Non-small cell lung, renal and thyroid cancers

Table 17.14 Examples of targeted therapies used in cancer treatment.

Drugs that target molecules involved in regulating cell proliferation

Alectinib is an inhibitor of anaplastic lymphoma kinase (ALK). This tyrosine kinase receptor belongs to the insulin receptor superfamily. Mutations such as the fusion of *ALK* with *EML4* (echinoderm microtubule-associated protein-like 4) produce the variant *ALK-EML4*. The inhibition of ALK by alectinib disrupts the ALK-mediated signalling pathway and eventually inhibits tumour cell growth, proliferation and resistance to

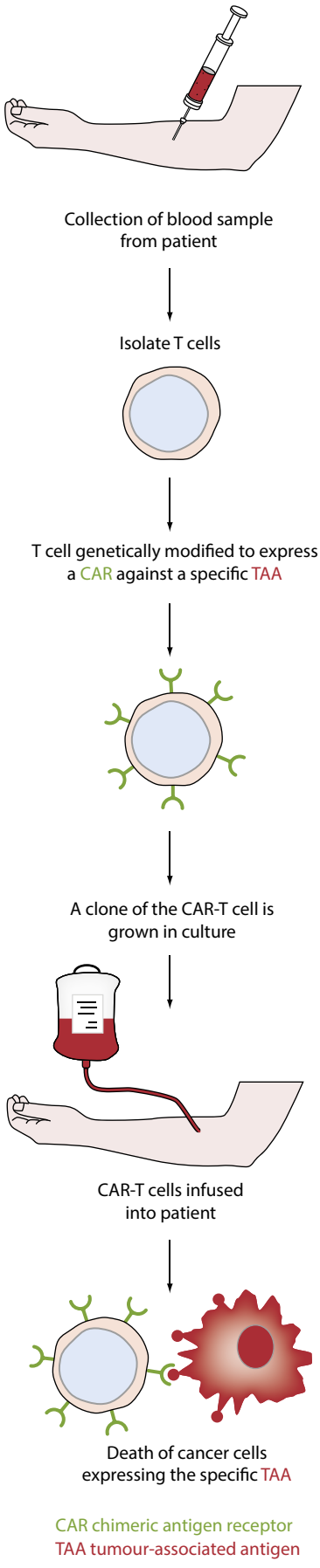

Collection of blood sample from patient
Isolate T cells
T cell genetically modified to express a CAR against a specific TAA
A clone of the CAR-T cell is grown in culture
CAR-T cells infused into patient
Death of cancer cells expressing the specific TAA
CAR chimeric antigen receptor
TAA tumour-associated antigen

Figure 17.38 Schematic illustration of chimeric antigen receptor T cell (CAR-T) therapy. A sample of patient blood is collected and its T cells isolated and then genetically modified by inserting into their genome a CAR gene for a T cell receptor able to recognise and bind to a specific tumour-associated antigen (TAA) expressed on the surface of the cancer cells.

apoptosis. Alectinib is used in the treatment of NSCLC expressing the ALK-EML4 fusion protein. Alpelisib is a phosphatidylinositol 3-kinase (PI3K) inhibitor, which inhibits PI3K activation in the PI3K/AKT kinase signalling pathway and also results in the inhibition of tumour cell growth and survival. Alpelisib is used to treat ER positive, epidermal growth factor receptor 2 (HER2) negative and PIK3CA-mutated, metastatic breast cancer in postmenopausal women and in male patients.

Cetuximab is a recombinant MAb that competitively binds to the extracellular domain of EGFR, thereby inhibiting the binding of epidermal growth factor. This is found in both normal and tumour cells but in the latter is often mutated or overexpressed, where it acts as a driver of tumorigenesis. The overexpression of EGFR has been linked with advanced cancers and a poor patient prognosis. Cetuximab prevents the activation of EGFR, resulting in the inhibition of signal transduction, and has an antiproliferative effect. Cetuximab is used for the treatment of metastatic colon, head and neck cancers that express the wild-type form of *RAS*. Patients expressing variants of *RAS* do not respond to cetuximab therapy because mutations in RAS lead to the constant activation of downstream signalling pathways. This promotes the growth and survival of cancer cells in a manner that is independent of EGF/EGFR stimulation.

Dabrafenib is an inhibitor of mutated forms of *BRAF*, which belongs to the family of serine/threonine protein kinases. These have roles in regulating the mitogen-activated protein kinase (MAPK)/extracellular signal-regulated kinase (ERK) signalling pathways (MAPKs are ERKs). These, in turn, are involved in a wide range of signal transduction cascades. Mutations in *BRAF* constitutively activate growth signals, leading to the proliferation of tumour cells. Dabrafenib is used alone or in combination with other drugs in the treatment of advanced malignant melanoma.

Drugs that target molecules that regulate apoptosis

Bortezomib, a specific inhibitor of the 26S proteasome that degrades ubiquitinated proteins, is a prominent inducer of apoptosis. Its action blocks proteins targeted for proteolysis by the proteasomes (*Margin Note 17.4*), which disrupts various cell signalling pathways, leading to cell cycle arrest and apoptosis. Bortezomib is used to treat multiple myeloma and mantle cell lymphoma. Venetoclax (see *Figure 17.9*) is an inhibitor of the anti-apoptotic protein BCL-2 (see *Box 17.3*), which is overexpressed in some cancers and is linked to increased drug resistance and the survival of tumour cells. Venetoclax mimics the actions of apoptosis-promoting proteins in that it binds to BCL-2 and stimulates apoptosis in tumour cells. It is used for treatment of acute myeloblastic and chronic lymphocytic leukaemias.

Drugs that target DNA repair molecules

Genomic instability, characterised by an accumulation of mutations, is a hallmark of cancer cells. However, the same genomic alterations that predispose a cell to become cancerous may also render it susceptible to targeted therapies. Olaparib and talazoparib are inhibitors of the nuclear sited poly(ADP-ribose) polymerase (PARP). Both bind to PARP and block PARP-mediated base excision repair and DNA single strand repair pathways. Thus, the cells become more reliant on alternative repair mechanisms, such as homologous recombination repair (HRR), which is dependent on BRCA proteins (*Section 17.4*). However, cells with mutated *BRCA*s are deficient in HRR proteins and cannot effectively repair damaged DNA. The accumulation of damaged DNA and its associated cellular stress can trigger apoptosis in the tumour cell. Inhibitors of PARP used in combination with chemotherapy or radiotherapy can have a synergistic effect with increased levels of cancer cell death. Olaparib and talazoparib are used to treat breast, ovarian and prostate cancers with *BRCA1/2* mutations (*Section 17.4*).

Drugs that target molecules that regulate angiogenesis

Angiogenesis is the formation of new blood vessels. Anti-angiogenic cancer therapy uses drugs that inhibit the development of blood vessels to a tumour. They thus inhibit tumour growth and metastasis by limiting its blood supply rather than targeting the tumour cells directly.

Angiogenesis is a complex biological process regulated by various pro- and anti-angiogenic molecules. Among these, the vascular endothelial growth factors (VEGFs) and their receptors (VEGFRs) play crucial roles via their associated signalling pathways. The VEGF family includes VEGF-A, -B, -C, -D and -E, and placental growth factor (PGF) recognises three different tyrosine kinase receptors: VEGFR1, 2 and 3. The VEGFs are highly expressed in different types of tumours. Their binding to receptors can stimulate the MAPK and PI3K signal pathways and lead to angiogenesis. Most anti-angiogenic drugs target the VEGF signalling system, including its ligands, receptors and molecules of the downstream pathways.

Margin Note 17.12
Folkman: a pioneer in the study of tumour angiogenesis

In 1971, Judah Folkman (1933–2008) proposed that tumour growth was angiogenesis dependent and that the inhibition of angiogenesis could be therapeutic in cancer care. Since then, anti-angiogenesis therapy has gained considerable attention. In 2004, the first anti-angiogenic drug, bevacizumab, was approved for the treatment of metastatic colon cancer.

Bevacizumab is a humanised MAb that targets VEGF-A and inhibits angiogenesis. Its use initiated the era of anti-angiogenesis therapy. In combination with chemotherapy, bevacizumab is used in the treatment of cervical, colorectal, fallopian tube, kidney, liver, NSCLC, ovarian and peritoneal cancers, and glioblastoma. Ramucirumab, also a humanised MAb, targets VEGFR2, which is expressed on endothelial cells and prevents binding of its ligands, VEGF-A, -C and -D. This inhibits the events associated with the signalling pathway; hence, it blocks the VEGF-stimulated proliferation, permeability and migration of endothelial cells. Ramucirumab is used as a single agent or in combination with chemotherapeutic drugs in the therapy of advanced colorectal cancer, gastro-oesophageal junction adenocarcinoma, stomach adenocarcinoma and NSCLC. Lenvatinib inhibits all three VEGFRs. Lenvatinib also inhibits other receptor tyrosine kinases, including FGFR1, 2, 3 and 4, the PDGF receptor α (PDGFRα), the receptor tyrosine kinase c-Kit, and activities of the oncoprotein that is rather bizarrely called rearranged during transfection (RET). These receptors play a central role in the activation of signal transduction pathways involved in regulating cell proliferation, migration and differentiation, and in apoptosis and angiogenesis. Lenvatinib is used for the treatment of patients with metastatic, radioactive iodine (RAI)-refractory thyroid cancer and kidney, liver and endometrial cancers.

Drugs that target epigenetic regulators

Epigenetics is usually defined as heritable changes in gene expression that are not attributable to alterations in the sequence of genomic DNA.

Vorinostat is used to treat cutaneous T cell lymphoma. It binds to the catalytic domain of histone deacetylases (HDACs), thereby inhibiting deacetylation and leading to an accumulation of hyperacetylated histones and transcription factors. Hyperacetylation of histone proteins results in the upregulation of the cyclin-dependent kinase p21, followed by arrest of the cell cycle at G1. Hyperacetylation of nonhistone proteins such as p53, α tubulin and heat-shock protein 90 produces additional antiproliferative effects. Belinostat also inhibits HDAC activity. It was developed as an orphan drug to treat peripheral T cell lymphoma after other cancer treatments did not work or had ceased to be effective therapies.

Tazemetostat is an inhibitor of EZH2, a subunit of zeste 2 polycomb repressive complex, which is encoded by *EZH2* sited at 7q36.1. The complex is a histone methyl transferase that is overexpressed or mutated in a variety of cancer cells and plays a key role in tumour cell proliferation. Tazemetostat inhibits the activity of both wild-type and mutated forms of EZH2 by specifically preventing the methylation of histone H3 lysine 27 (H3K27). This results in decreased tumour cell proliferation in EZH2 mutated cancer cells. Tazemetostat is used to treat metastatic epithelioid sarcoma or if the tumour cannot be removed with surgery.

Gene therapy

Cancer gene therapy involves the delivery of specific genetic material into tumour cells to modify the expression of a gene, induce tumour cell death or slow the growth of the tumour. Cancer gene therapy strategies include:

- repairing or replacing missing or nonfunctioning tumour suppressor genes
- blocking the expression of an oncogene by using an antisense (RNA/DNA) approach
- expressing a gene to induce apoptosis or enhance tumour sensitivity to conventional chemotherapy or radiation therapy
- enhancing the immunogenicity of the tumour to stimulate immune system defence

In a sense, HSCT is a form of gene therapy. However, the term, also called genome editing, is usually used to describe a group of techniques that can insert, delete or replace specific base sequences in genomic DNA. Gene editing has broad applications in research and clinical practice, including treating genetic disorders and cancers.

The three main gene editing tools include are:

1. zinc finger nucleases
2. transcription activator-like effector nucleases
3. RNA-guided clustered regularly interspaced short palindromic repeats (CRISPR)–Cas nuclease systems

The RNA-guided CRISPR–Cas nuclease systems are the most widely used genome editing technology because of their simplicity of design, high efficiency and precision, and are less time consuming to implement. Repetitive DNA sequences (CRISPR) were

Margin Note 17.13
Epigenetic alterations and cancer

Epigenetic modifications are normal and essential biological processes. Some common epigenetic modifications include DNA methylation and histone modifications (such as acetylation, methylation and phosphorylation). Epigenetic deregulation often includes mutations and/or the aberrant expression of chromatin-modifying enzymes, their associated proteins or abnormal levels of noncoding RNA in cells. This can lead to inappropriate initiation or inhibition of gene expression that ultimately contributes to the development of a tumour. Cancer development at all stages is driven by both genetic alterations and epigenetic abnormalities. Treatments that target epigenetic abnormalities are becoming an attractive strategy in cancer therapy.

Margin Note 17.14
Gendicine

Gendicine was the first gene therapy product to gain clinical approval, in this case by the China Food and Drug Administration (CFDA). It introduces a functional *TP53* into cancer cells using a viral vector and is used to treat head and neck cancers. Chimeric antigen receptor T cell (CAR-T) therapy is an example of gene editing to enhance tumour immunogenicity.

identified in prokaryotes in 1993. However, it was not until 2007 that their function in prokaryotic immunity was discovered. Prokaryotic cells are able to capture small fragments of DNA from invading viruses and use them to form bacterial DNA segments known as CRISPR arrays. The CRISPR arrays form a memory bank in the prokaryotic cell. If the prokaryote is infected by the same or a closely related virus, it is able to synthesise RNA segments using the CRISPR array as a template. The RNA segments can recognise and target the viral DNA and direct a Cas nuclease to hydrolyse it, thus eliminating the viral infection.

The CRISPR–Cas gene editing method was developed from the naturally occurring genome editing system of prokaryotic cells. A small piece of RNA (called guide (g) RNA) is synthesised, which is composed of two parts: (1) CRISPR RNA (crRNA), a 17–20 nucleotide sequence complementary to the target DNA and (2) a transactivating CRISPR RNA (tracrRNA) that serves as a binding scaffold for the Cas nuclease. In nature, the gRNA is a duplex molecule, with crRNA and tracrRNA annealed together. Synthetically, single guide RNA (sgRNA) can be produced with these two molecules connected by a linker loop. The sgRNA and Cas form a ribonucleoprotein (RNP) complex (*Figure 17.39*). As in prokaryotes, the sgRNA recognises the target DNA sequence adjacent to a protospacer adjacent motif (PAM). When the RNP binds to the target DNA, the Cas facilitates the unwinding of the DNA double strands at the target site and hydrolyses the DNA at 3′ base pairs upstream of the PAM region. In this way, a mutated sequence in a gene can be removed. The normal DNA repair machinery of the cell can then insert a customised DNA sequence into the genome, with the result that the wild-type genome is restored and can be expressed.

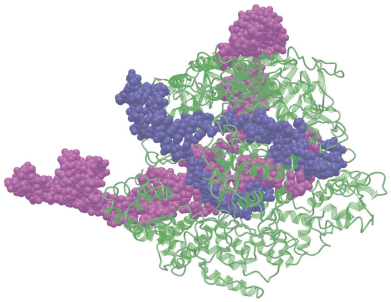


Figure 17.39 Molecular model of a *Streptococcus pyogenes* CRISPR–Cas9 complex with a guiding single-stranded RNA (magenta), double-stranded target DNA (blue) primed for cleavage and a Cas9 nuclease (green). PDB file 5F9R. ⏎

BOX 17.8 ARTIFICIAL INTELLIGENCE IN ONCOLOGY ⏎

Artificial intelligence (AI) was defined by John McCarthy (1927–2011) in 1956 as 'the science and engineering of making intelligent machines'. It is a branch of computer science that seemingly mimics human intelligence to perform specified tasks. In general terms, it relies upon the ability of computer algorithms to learn and recognise patterns and interactions from a large set of appropriate previously defined training examples, each with its desired output or classification. Artificial intelligence systems detect patterns and features in such examples, which allows them to extrapolate from these data and identify new, that is, undefined, samples, and also to make suitably useful predictions. For example, AI can identify new biomarkers within large sets of clinical data to assist in the detection and diagnosis of diseases and to improve clinical screening programmes, and, for instance, to make predictions, for example about the conformation of a protein using only its known primary structure.

In cancer biology, AI is used to improve the standard approaches to the clinical management of patients in detecting, diagnosing and staging tumours, and monitoring any changes within them over time. Artificial intelligence systems are particularly adept at analysing visual data. They therefore have major applications in analysing clinical images, for example chest X-ray pictures of potential lung cancer patients or computed tomography (CT) and magnetic resonance imaging (MRI; see *Chapter 14*) scans taken during routine cancer care. Artificial intelligence is also being used in the treatment of cancers, for example in contouring, which is an essential adjunct in radiotherapy treatment. Contouring involves outlining the target volumes and organs at risk (OAR) to maximise the effectiveness of the treatment and minimise radiation toxicity. Artificial intelligence systems, trained using CT or MRI scans, can improve the efficiency of contouring by automatically determining the OAR and sometimes the target volume prior to radiotherapy. Artificial intelligence may also improve personalised cancer treatment and clinical management. The precise treatment of any one tumour is essential for a beneficial patient outcome. Unfortunately, tumours are highly heterogeneous in their mutated genes, protein complement and cell phenotypes; therefore, the same surgery or chemotherapy may have different outcomes in patients suffering from the same type of tumour. The analyses of large sets of relevant data by AI systems can reveal clinically relevant patterns of information, not immediately obvious within the data, and lead to more suitable personalised treatments giving improved clinical outcomes for patients, while also leading to more accurate prognoses.

The AI systems approved for clinical use by the Food and Drug Administration (FDA) have been mainly of use in cancer diagnostics, with the largest impact on clinical practice in breast, lung and prostate oncology-related areas. In the UK, the NHS has been given regulatory approval by the National Institute for Health and Care Excellence (NICE) to use several AI systems to aid contouring in patients requiring radiotherapy. However, it should be noted that, currently, all AI systems in clinical use are intended as aids to and not as substitutes for generally accepted methods. It would, however, seem certain that AI has the promise of proving ever more useful in diagnosing, treating and the clinical management of cancer patients.

17.8 SPECIFIC TYPES OF CANCERS

Cancer patients, naturally, may present with a variety of symptoms depending on the origin and extent of their tumour. The treatment administered for any cancer patient therefore depends largely on the type of cancer, its location, the stage or grade of the

disease, and the age and otherwise general health of the patient. Over 200 different types of cancer have been described; among the most likely to be encountered clinically are:

- breast
- colorectal
- lung
- prostate
- bone marrow (leukaemias)

Breast cancer

Breast cancer is a malignant neoplastic disease of breast tissues. They can develop in the milk-producing tissues or in the ducts of the glands (*Figure 17.40*). Ductal *in situ* carcinoma (DCIS) is a malignant growth confined within the system of ducts. If diagnosed at this stage, the disease is almost completely curable. Patients with the disease, however, have an 80–90% increased risk of developing an invasive carcinoma. Breast cancer is the most common cancer: the incidence of new cases in the UK is about 56 000 annually. In the USA, it is roughly 240 000. The annual incidence worldwide is approximately 2.3 million, causing about 670 000 deaths. In the UK and USA, breast cancer is the second leading cause of cancer deaths in women. In the UK, nearly 11 500 women die of breast cancer each year. It is thought that about one in seven women will develop breast cancer. Breast cancer also occurs in men. In the UK, approximately 375 men are diagnosed yearly, meaning the incidence for men is 1 in 870. Approximately 85 men die annually in England of the disease.

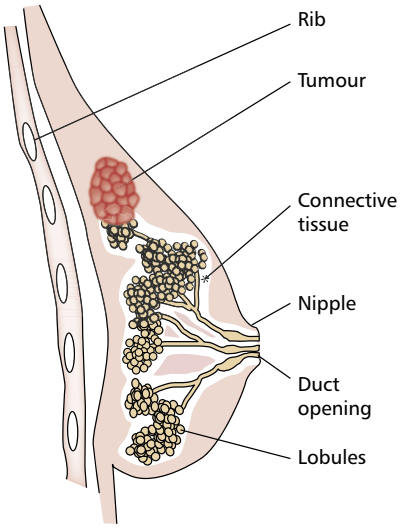


Figure 17.40 Schematic illustrating the presence of a cancer of the breast.

Risk factors for breast cancer

It is apparent from the above that the greatest risk factor is being female; the next is increased age. Several hormonal factors increase the risk of breast cancer developing, for example an early menarche or late menopause, childlessness or having a first child later in life. The long-term (greater than five years) use of oestrogen/progesterone hormone replacement therapy gives a small increased risk, as does the use, particularly of older types, of contraceptive pills. Other risk factors include being overweight or obese and the regular consumption of alcohol and a familial history of the disease. Variants of many genes are linked to the development of breast cancer; about 5% of all of them are related to the inheritance of mutated *BRCA1* and *BRCA2*, as described in *Section 17.4*.

The much lower risk of men developing breast cancer is because the male breast lacks terminal duct lobular units or lobules: the epithelial structures that produce milk and the major site of breast cancer precursors. Also, the male breasts are much less sensitive to hormones than female ones (see below).

Signs and symptoms of breast cancer

Symptoms of breast cancer include a new painless lump occurring in a breast, changes in the size or shape of the breast or in the position of the nipple, a retracted nipple or a discharge from it, an eczematous rash around the nipple or a thickening behind the nipple, puckering or dimpling of the skin of the breast or a swelling or lump in the armpit. Most breast cancers are detected by a mammogram, which uses X-rays to locate the position of the potential tumour. Mammography is also used to screen for breast cancer and, in the UK, this is offered every three years to all women over the age of 50 years. The lump may also be examined by ultrasound, which gives a picture of the blood supply to the lump. Microscopic examination of cells from a fine needle aspirate of the lump may also be helpful. Alternatively, the lump may be excised under a general anaesthetic for pathological examination. A microscopic examination of cancer cells is useful in determining treatment. Less aggressive cancer cells appear to resemble differentiated normal cells, whereas those of abnormal appearance are characteristic of fast growing and aggressive cancers.

Diagnosis, staging and treatment of breast cancer

A diagnosis of breast cancer requires tumour cells to be examined immunohistochemically or by FISH (*Section 17.6*) to detect the expression of receptors for oestrogen, progesterone and HER2 (also called ErbB2). Treatment is largely determined by the hormone status of the tumour and its stage (see below). Breast cancer cells expressing receptors for oestrogen are referred to as ER positive (ER+). Those with low levels or a lack of ERs are referred to as ER negative (ER–). About 80% of breast cancers are ER+. Breast cancers may also have receptors for progesterone (PR+). The aim of hormone therapy is to prevent oestrogen from stimulating cancer cell proliferation; therefore, it is most effective against ER+ breast cancer. Tamoxifen, one of the most used hormone therapies, competes with

oestrogen for its receptor (*Section 17.7*). It is suitable for use with both premenopausal and postmenopausal women, and men with breast cancer. Fulvestrant is a type of hormone therapy for postmenopausal women with secondary breast cancer that is hormone receptor positive. It may be used if tamoxifen or aromatase inhibitors are no longer controlling the cancer (*Margin Note 17.15*). Fulvestrant works in two ways. It blocks ERs in a similar manner to tamoxifen, but also denatures the receptors, so they are not recognised by oestrogen. Again, it may be in combination with other drug regimens.

In the UK, all women with early-stage breast cancer are tested for the HER2 status of their tumour. Epidermal growth factor 2 is overexpressed in approximately 30% of human breast cancers, which disrupts the normal control mechanisms governing cell growth and gives rise to aggressive tumour cells and increased breast cancer metastasis. Targeted cancer drugs prevent HER2 from binding to its receptor or inhibit intracellular signalling mechanisms, and so restrict cell division and growth. Hence, assessing the HER2 status of breast cancer cells can indicate whether targeted drugs therapy is indicated. Drugs used in targeted therapy include lapatinib, neratinib and tucatinib. The humanised MAb trastuzumab (perhaps better known as Herceptin) also targets HER2 and can be used in combination with chemotherapy. A second humanised MAb, pertuzumab, blocks the actions of HER2 and is used in combination therapy with trastuzumab and chemotherapy in patients presenting with advanced HER2+ breast cancer.

About 15% of breast cancers are triple negative, that is, they lack receptors for oestrogen, progesterone and HER2. This condition is more common in younger women and, unfortunately, hormone therapies and HER2-targeted cancer drugs are not effective against this type of breast cancer. Chemotherapy is therefore the preferred treatment.

A simplified TNM system for staging of breast cancer is given in *Table 17.15*. Knowing the stage of the cancer allows therapy and care of the patient be optimised.

	Definition
T (tumour)	
TX	Primary tumour cannot be assessed
Tis	Ductal carcinoma *in situ* (DCIS), a preinvasive breast cancer
T1	Tumour is 20 mm or less across
T2	Tumour is more than 20 mm but less than 50 mm in size
T3	Tumour is bigger than 50 mm
T4	T4a: tumour has spread into the chest wall T4b: tumour has spread into the skin and the breast may be swollen T4c: tumour has spread to both the skin and the chest wall T4d: inflammatory carcinoma
N (node)	
NX	Regional lymph nodes cannot be assessed
N0	No tumour cells in any nearby nodes
N1	N1mi: micrometastases, tumour size in nearby lymph nodes is 0.2–2 mm N1a: tumour cells in one to three lymph nodes in the armpit and at least one is larger than 2 mm N1b: tumour cells in the sentinel lymph nodes behind the breastbone (the internal mammary sentinel nodes) N1c: tumour cells in one to three lymph nodes in the armpit and in the sentinel lymph nodes behind the breastbone
N2	N2a: tumour cells in four to nine lymph nodes in the armpit and at least one is larger than 2 mm N2b: tumour cells in the lymph nodes behind the breastbone, which can be seen on a scan or felt by a clinician. There is no evidence of cancer in the lymph nodes in the armpit
N3	N3a: tumour cells in 10 or more lymph nodes in the armpit and at least one tumour is larger than 2 mm or there are cancer cells in the nodes below the collarbone N3b: tumour cells in lymph nodes in the armpit and lymph nodes behind the breastbone N3c: tumour cells in lymph nodes above the collarbone
M (metastasis)	
M0	Tumour metastasis not apparent
M1	Detectable, metastatic tumours larger than 0.2 mm

Table 17.15 Simplified TNM staging for breast cancer. ⏎

The first line of treatment for breast cancers is mastectomy, that is, the partial or complete surgical removal of one or both breasts. A partial mastectomy may involve excision of the lump and some of the surrounding tissue. Chemotherapy or hormonal therapy may be given to reduce the size of the tumour prior to excision. During surgery, lymph nodes are also removed from the armpit and used for accurate staging of the tumour. Two to four weeks after excision or mastectomy, radiation therapy is used to destroy any remaining cancer cells. Surgery for DCIS may be followed by radiation or hormone therapy with tamoxifen (*Figure 17.41*), in most cases to prevent a possible progression. A complete mastectomy is likely to be a traumatic event for the patient, particularly females,

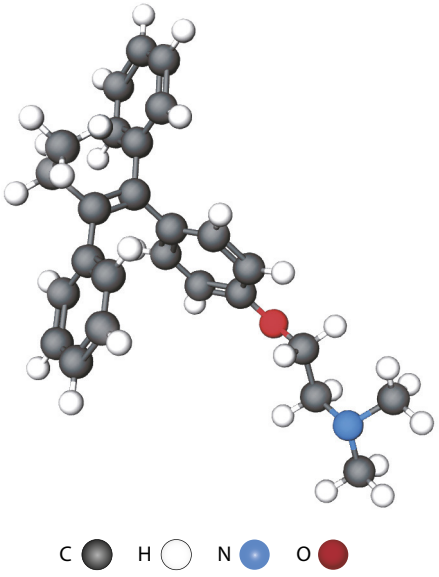


Figure 17.41 Computer drawn model of tamoxifen. ⏎

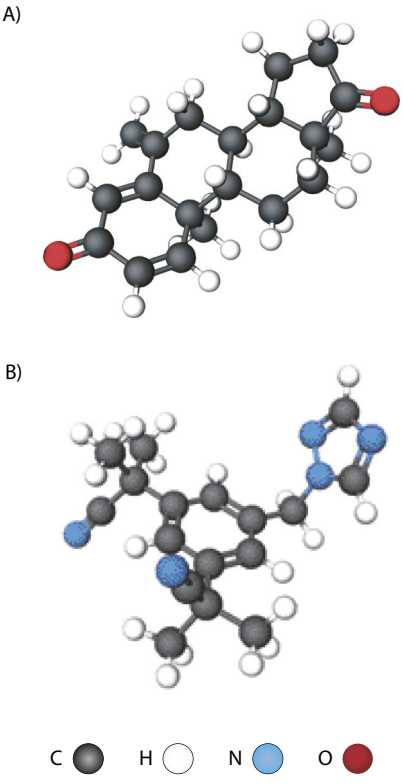


Figure 17.42 Computer drawn models of the anticancer drugs (A) exemestane, a steroid (compare to *Figure 7.68*), and (B) the nonsteroidal anastrozole. ⏎

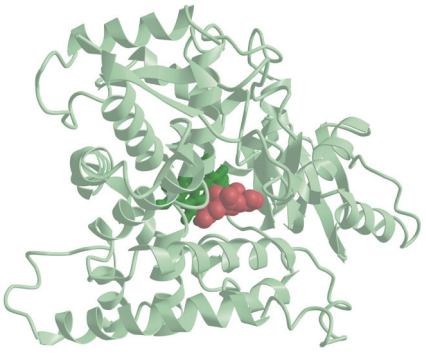


Figure 17.43 Molecular model of cytochrome P-450 aromatase (light green) complexed with the anticancer drug exemestane (red). The haem coenzyme is indicated in dark green. PDB file 3S7S. ⏎

Margin Note 17.15 ⏎
Aromatase and anastrozole

The oestrogen receptor is expressed in up to 80% of breast cancers. The binding of oestrogen to the receptor stimulates tumour cells to proliferate. The final step in the biosynthesis of oestrogens is their formation from precursor androgens in a reaction catalysed by cytochrome P-450 aromatase (CYP19A1), also called oestrogen synthase:

Androgen
+
$3NADPH + 3H^+ + 3O_2$
↓ Cytochrome P-450 aromatase
oestrogen
+
$3NADP + 3H_2O + HCOOH$

Most mixed function oxidases (monooxygenases) have broad substrate specificities (*Chapter 12*); however, aromatase catalyses the conversion of androgens such as testosterone, androstenedione and 16α-hydroxytestosterone to oestrone, 17β-oestradiol and 17β,16α-oestriol with a high specificity. Thus, aromatase is a key target for anticancer therapy. For example, the steroid exemestane and nonsteroidal agents, such as anastrozole (*Figure 17.42*), are used to treat postmenopausal oestrogen-dependent (receptor-positive) breast cancer. These drugs bind to aromatase (*Figure 17.43*) and inhibit its activity. This effectively reduces the concentration of circulating oestrogen by preventing its formation from the low levels of circulating androgens and therefore reduces tumour development.

Anastrozole has been used to treat breast cancer for many years. However, in 2023, it received wide publicity when NHS England approved its use as a cancer prophylaxis in postmenopausal women at moderate or high risk of developing the disease. It is thought that if only a quarter of the estimated 289 000 women at risk used anastrozole, then about 2000 cases of breast cancer would be prevented in England and, much less importantly, this would save the NHS about £15 million in treatment costs.

and must be dealt with in a sensitive manner. Whatever the type, surgery is always combined with radiation therapy and hormonal and/or chemotherapies, with the exact combinations depending on the hormonal status of the patient. Patients with a hormone receptor positive tumour may benefit from treatment with inhibitors of aromatase, such as anastrozole (*Margin Note 17.15*).

The prognosis for breast cancer patients has greatly improved since the early 1970s, given that mortality rates have decreased by roughly 40%. Approximately 76% of women diagnosed with breast cancer in England live for 10 or more years. However, the rate of mortality varies with age. About 81% of women 15–44 years old survive their disease for 10 years. The 10-year survival for those aged 75–99 years is approximately 57%. The value at roughly 87% is highest in the 55- to 64-year-old bracket, because younger patients tend to suffer more aggressive neoplasms and older women are less able to combat the disease or cope with the therapy. However, patients who are HER2+ have substantially lower overall survival rates and shorter disease-free intervals than HER2– sufferers.

Colorectal cancer

Colorectal cancer (CRC), also known as bowel cancer, is a malignant neoplasm of the large intestine. It is one of the three most common cancers in men and women, in both the USA and the UK. In the UK, there are nearly 43 000 new cases yearly, which forms about 11% of all cases of cancer. The majority occur in men, but CRC is the third most common cause of cancer deaths in men and women, killing approximately 16 800 annually. In the USA, CRC is the second most common cause of cancer deaths. In 2024, it was estimated that approximately 153 000 patients will be diagnosed with it and over 53 000 will die from the disease. However, CRC is a global problem. The WHO has indicated that, worldwide, CRC is the second leading cause of cancer-related deaths. In 2020, it was estimated there were more than 1.9 million new cases and over 930 000 deaths. However, large geographical variations in incidence and mortality were observed: incidence was highest in Europe, Australia and New Zealand, and the mortality was greatest in eastern Europe. *Table 17.16* lists the different types of CRCs, but the vast majority are adenocarcinomas and this section will concentrate on them.

Colorectal cancer	Description
Adenocarcinoma	Constitute 95% of colorectal cancers; they arise from mucus-secreting cells in the epithelium of the GIT lining; most produce mucin and 1–2% are signet ring tumours where the intracellular mucus has pushed the nucleus to one side
Squamous cell carcinoma	Arise from epithelial cells lining the GIT
Carcinoid tumour	Rare, slow-growing tumour of neuroendocrine origin
Sarcoma	Majority are leiomyosarcomas arising from smooth muscle tissue in the GIT wall
Lymphomas	Form 1% of colorectal cancers; originate from lymphoid cells in the GIT wall

GIT, gastrointestinal tract.

Table 17.16 Types of colorectal cancer.

Risk factors for colorectal cancer

The risk factors for CRC include increasing age, a diet rich in fats and dairy products and low in fibre, a history of inflammatory bowel disease, and a hereditary predisposition given a familial history of bowel cancer. Adenocarcinomas are known to originate from preexisting polyps. Indeed, the sequence of development of a polyp to a neoplasm is well documented. Polyps are benign tumours that are mostly asymptomatic. They arise from the epithelium of the GIT and range in size from a small bump to a lesion 30 mm in diameter. The development of CRC is also associated with both hypomethylation and hypermethylation of DNA and this can be detected in polyps. Methylation of DNA is involved in switching genes off; hence, hypomethylation can result in activation of oncogenes, whereas hypermethylation could lead to inactivation of tumour suppressor genes (*Margin Note 17.13*).

Colorectal cancers result from the accumulation of several distinct genetic variations. A deletion in the short arm of chromosome 17 leads to an inactive tumour suppressor

p53. Deletions in the long arms of chromosomes 12 and 18 mean that activity of *KRAS*, a protooncogene sited at 12p12.1, is impaired and the tumour suppressor gene *DCC*, normally located at 18q21.2, is deleted. Indeed, allelic deletions involving chromosome 18 are found in more than 70% of CRCs. Mutations in the oncogene *RAS* have been identified in about 50% of patients diagnosed with metastatic CRC.

Hereditary nonpolyposis colorectal cancer (HNPCC) or Lynch syndrome occurs in several forms that are associated with inherited mutations in a number of genes. The condition is estimated to cause 2–5% of CRCs and is linked to 40% of cases of CRC in people less than 30 years old. Indeed, HNPCC is the most common cause of inherited colorectal cancer. Individuals with adenomatous polyposis familial (FAP), a rare condition responsible for 1% of colorectal cancers, have a mutated form of *APC*. The product of this tumour suppressor gene is a multidomain protein (APC) that antagonises the wingless-related integration site (WNT; see *Table 17.4*) signalling pathway and also degrades β-catenin. Mutations that result in the loss of these functions contribute to cancer progression by an inappropriate activation of the WNT pathway and the production of an inactive β-catenin unable to inactivate growth-promoting oncogenes, such as c-*MYC*. Patients with FAP are almost certain to develop CRC in middle age and are recommended to have a prophylaxis removal of their colon by the age of 25 years.

Additional risk factors for CRC include diabetes mellitus (*Chapter 7*), Crohn disease (*Chapter 11*) and being of Ashkenazi Jewish origin. Members of this ethnic group also have a higher incidence of mutated *BRCA1* and *BRCA2* (*Section 17.4*) and are, therefore, also more likely to develop breast cancer. The genetic link between these genes and CRC has not been fully elucidated.

Signs, symptoms and diagnosis of colorectal cancers

Most CRC patients are diagnosed aged 50–60 years, with the incidence peaking at 85-plus years. Many early-stage CRCs are detected by bowel cancer screening programmes (*Margin Note 17.16*).

Patients may present with a loss of weight, bleeding from the rectum and/or blood in the faeces, anaemia resulting in fatigue and breathlessness, changes in bowel habits including increased diarrhoea or constipation, and pain in the abdomen. An initial detection of a lump in the abdomen or rectum means that the patient must be referred for further specialist examination. Colonoscopy is used to examine the whole colon and to obtain a biopsy for further investigation. An alternative to colonoscopy is flexible sigmoidoscopy, an endoscopic technique that examines only the lower part of the colon and the rectum. The concentration of CEA (*Table 17.8*) in the blood can be measured, although increased values are not entirely specific for cancer and a poorly differentiated tumour may not produce CEA. This measurement may, however, be useful for monitoring disease progress and treatment. Abdominal CT scans are helpful in diagnosing metastatic spread to the lymph nodes and liver, and chest X-rays are routinely used to detect metastases in the lungs.

Staging and treatment of colorectal cancer

Several staging systems have been used for colorectal cancer. The Dukes previously used has been largely superseded by a TNM staging system (*Table 17.17* and *Figure 17.45*).

The usual treatment for CRC is the surgical removal of the affected parts with resection (see *Figure 11.49*). The advantage of resection is that the patient does not require a colostomy bag. Patients with rectal cancer may be given radiation treatment following surgery to reduce the risk of the tumour recurring locally. Patients successfully treated by surgery should be regularly checked for several years to detect any recurrence of the tumour. Radiation therapy may also be used in the palliative care of late-stage metastatic cancer.

Chemotherapy is applicable if the patient is unsuitable for surgery, after surgery to try and stop remission or if the cancer has metastasised. Chemotherapy usually involves combination therapy involving two or three drugs. The most common drugs for treating CRCs are capecitabine; 5FU; the combination of folinic acid, 5FU and oxaliplatin (*Section 17.6*) or irinotecan; and combined oxaliplatin and capecitabine. Metastatic CRCs, sometimes referred to as advanced bowel cancer, are treated with the human MAb, bevacizumab and panitumumab, and the chimeric MAb cetuximab (*Section 17.6*).

Prognosis depends on the age and otherwise general health of the patient, but largely on the stage at which the cancer was detected. Following diagnosis, approximately 90% and 85% of stage 1 and 2 patients will survive for five years or more, respectively. This falls to 65% for stage 3. If the cancer has spread to the liver and that organ can be removed surgically, more than 40% will survive for five or more years after their operation. Unfortunately, only about 10% of stage 4 patients survive five years or more after their diagnosis. The prognosis is bleak for patients with metastases of the liver and lungs.

Margin Note 17.16 ⏎ The NHS bowel (colorectal) cancer screening programme

Initially, the NHS bowel (colorectal) cancer screening programme was available only to those aged 60–74 years, although adults over 74 can request to be kept on the programme. However, in 2021, the programme began being expanded, over the course of four years, to include those aged 50–59 years. Everyone of suitable age, registered with a general practitioner and living in England is automatically sent an NHS bowel cancer screening faecal immunochemical test kit (*Figure 17.44*) every two years. A small sample of faeces is collected on a small plastic stick by the individual and placed into the sample bottle provided. This is then posted cost free to a laboratory to be tested for occult blood. If blood is not found in the sample, then another screening test will be necessary in two years. Note, this result does mean bowel cancer is necessarily absent, and any of the other signs and symptoms of its presence should be investigated immediately. About 98% of people do not need further tests, but if blood is detected, it could indicate the presence of polyps or bowel cancer. However, it does not necessarily indicate either, given that other conditions, for example haemorrhoids, would result in blood in the faeces. Individuals with a positive test for occult blood are offered further tests, such as a colonoscopy, to investigate its reason.

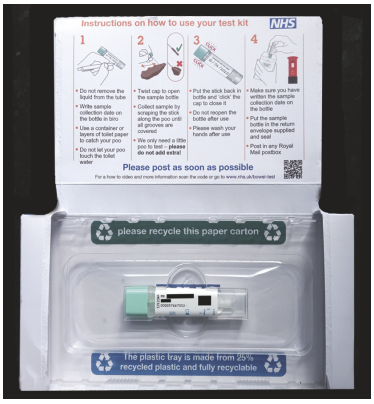

Instructions on how to use your test kit
NHS
Please post as soon as possible
please recycle this paper carton
The plastic tray is made from 25% recycled plastic and fully recyclable

Figure 17.44 A bowel cancer testing kit. Courtesy of the NHS bowel cancer screening programme. ⏎

Definition	
T (tumour)	
Tis	Carcinoma *in situ* or intramucosal carcinoma
T1[a]	Tumour has penetrated the muscularis mucosa into the submucosa
T2	Tumour has grown into the muscularis propria
T3	Tumour has grown into the outermost layers of the colon or rectum but has not penetrated them
T4	T4a: tumour has penetrated the wall of the colon or rectum but not grown into nearby tissues or organs T4b: tumour has grown through the wall of the colon or rectum and is attached to or has grown into neighbouring tissues or organs
N (node)	
N1	Tumour has spread to one to three nearby lymph nodes or into areas of fat near the lymph nodes
N2	N2a: tumour has spread to four to six nearby lymph nodes N2b: tumour has spread to seven or more nearby lymph nodes
M (metastasis)	
M0	Tumour metastasis absent
M1	M1a: tumour has spread to one distant organ or set of lymph nodes but not to distant parts of the peritoneum M1b: tumour has spread to more than one distant organ or distant set of lymph nodes but not to distant parts of the peritoneum M1c: tumour has spread to distant parts of the peritoneum and may/may not have spread to distant organs or lymph nodes

[a] Tumour stages T1, T2 and T4 are equivalent to Dukes A, B and D, respectively. See also *Figure 17.45*.

Table 17.17 A simplified description of the TNM staging system for colorectal cancer.

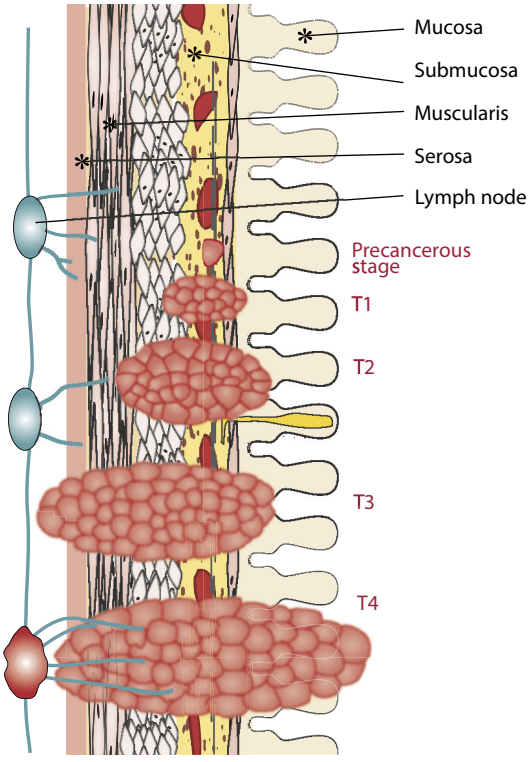


Figure 17.45 Schematic illustration of the colorectal cancer TNM tumour stages 1–4.

Lung cancer

Lung cancer is a pulmonary malignancy (*Figure 17.46*). Worldwide, lung cancer is the leading cause of a cancer-related deaths and is the most common cancer in men and the second most common in women. In 2020, more than 2.2 million new cases of lung cancer occurred globally. In the UK, approximately 48 000 new cases of lung cancer are diagnosed annually, which is approximately 13% of all cancers. The disease kills over 34 000, more than 21% of all cancer deaths. In the USA, lung cancer is the third most common malignancy after prostate cancer in men and breast cancer in women, with over 234 000 new cases and 125 000 deaths thought to have occurred in 2024.

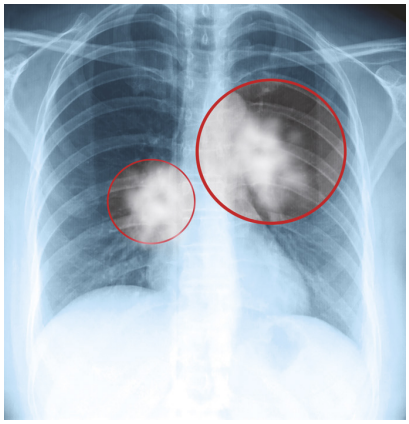


Figure 17.46 False colour X-ray of human lung cancers. See also *Figure 1.6*. Courtesy of Shutterstock.com/muratart.

Risk factors for lung cancer

The risk of lung cancer increases with age, the incidence peaking at 85-plus years in the UK. Most lung cancers are caused by tobacco abuse, hence the decreasing incidence seen in men but the increasing pattern in women, which reflects the historical patterns of smoking (see *Box 17.6*). In the UK, over 70% of lung cancer cases and more than 85% of lung cancer deaths are linked to carcinogens in cigarette smoke. These include benzo[a]pyrene (a polycyclic hydrocarbon) and *N*-nitrosamines, both of which are metabolised by the cytochrome P-450 enzymes in the liver to carcinogens that form adducts with DNA (*Chapter 12*). Exposure to occupational carcinogens is thought to cause approximately 13% of all cases of lung cancer. Carcinogens encountered in this manner include asbestos, silica and industrial solvents. Air pollution, especially diesel fumes, is linked to about 8% of cases, and particularly affects professional drivers. Radon, a naturally occurring radioactive gas produced from the decay of uranium, is found at relatively high levels inside homes built of, and on, granite. The gas is responsible for a small but statistically significant number of cases each year. People who work in uranium mines are also at risk of radon-induced lung cancer.

Classification of lung cancers

Lung cancers are classified into two main groups:

1. small cell lung cancer
2. non-small cell lung cancer

Small cell lung carcinoma (SCLC) accounts for approximately 15–20% of all lung cancers. It is a neuroendocrine carcinoma. The cancer cells are relatively small, hence the name, with a round-to-fusiform shape, a high nucleus-to-cytoplasm ratio, finely granular nuclear chromatin and absent or inconspicuous nucleoli. Non-small cell lung cancer can be subdivided into several types, and three of the most common types are squamous cell carcinomas, adenocarcinomas and large cell carcinoma. Squamous cell carcinomas develop from epithelial cells lining the respiratory tract. Adenocarcinomas arise from mucus-secreting cells in the lining of the respiratory tract. Finally, large cell carcinomas, not surprisingly, consist of cells that are relatively large and rounded compared with the other forms. Less common types of NSCLC include adenosquamous carcinoma and sarcomatoid carcinoma.

Symptoms, diagnosis and staging of lung cancer

The symptoms of lung cancer at presentation include a persistent and nagging cough, shortness of breath, coughing up blood-containing sputum (**haemoptysis**), chest pains when breathing or coughing, an unexplained loss of weight and recurrent chest infections, such as pneumonia and bronchitis. A patient showing any of these symptoms should be referred urgently for a chest X-ray and/or a CT scan. Some patients are asymptomatic and may only be diagnosed following a routine chest X-ray. Lung cancer does not generally cause symptoms in its early stages. Indeed, often by the time they are diagnosed, the disease is generally in an advanced state. Other tests include bronchoscopy (*Figure 17.29*) and an image-guided fine needle biopsy, such as endobronchial ultrasound-guided transbronchial needle aspiration (EBUS-TBNA). Intrathoracic lymph node assessment is offered to people with suspected lung cancer who have enlarged intrathoracic lymph nodes (lymph nodes greater than or equal to 10 mm short axis on CT).

If diagnostic tests indicate lung cancer, the tumour should be staged by scanning patients using PET or PET-CT. A TNM staging system for lung cancer, akin to those mentioned for breast and colorectal cancers (*Tables 17.15* and *17.17*), has been devised by the UICC (*Table 17.18*). Staging of the disease is required to determine its treatment.

Stages				
0	I	II	III	IV
TisN0M0	Stage IA1 T1aN0M0 T1(mi)N0M0	Stage IIA T2bN0M0	Stage IIIA T(1–2)N2M0 T3N1M0 T4N(0–1)M0	Stage IVA Any T, any N, M1a,b
	Stage IA2 T1bN0M0	Stage IIB T(1–)N1M0 T3N0M0	Stage IIIB T(1–2)N3M0 T(3–4)N2M0	Stage IVB Any T, any N, M1c
	Stage IA3 T1cN0M0			
	Stage IB T2aN0M0			

Table 17.18 The NHS guidelines for the TNM staging of lung cancer as defined by the Union for International Cancer Control (UICC). ⏎

Treatment of lung cancer

Patients who abuse tobacco are advised to stop as soon as the diagnosis of lung cancer is suspected. For NSCLC, the standard of care for stages I, II and IIIA disease is surgery. How much of the lung is removed depends on the stage of disease and the health of the patient. During surgery, lymph nodes are sampled to aid more accurate staging of disease. Radiotherapy is recommended for patients with stage I–II disease who decline surgery or patients who are medically inoperable. Radiotherapy is also used with stage III patients who cannot tolerate or who decline chemoradiotherapy (with or without surgery). Some patients with stage II–III disease might have chemotherapy before and after surgery or radiotherapy treatment.

Patients with stage I–III SCLC are treated with a cisplatin-based combination chemotherapy. In those with impaired renal function, carboplatin can replace cisplatin for its lower risk of nephrotoxicity. The drug regimen may be given in cycles to patients who respond to treatment. Patients with limited stage (stage I–III) SCLC are given radiotherapy concurrently with chemotherapy. The aim for patients with stage IV lung cancer is to control the disease for as long as possible and help to ease symptoms. Depending on the patient's situation, one or more treatments of targeted therapy, immunotherapy, chemotherapy and radiotherapy might be applied. Patients might be asked if they wish to participate in a clinical trial as part of the treatment. This might be to test a new treatment or to look at different combinations of existing treatments.

Prostate cancer

Prostate cancer is a disease of the prostate: a gland that surrounds the first part of the male urethra, weighing 7–20 g and up to 25 mL in volume (*Figure 17.47*). It and the seminal vesicles help produce the fluid component of semen that supports the spermatozoa produced by the testes (*Chapter 7*; see *Figure 7.66*).

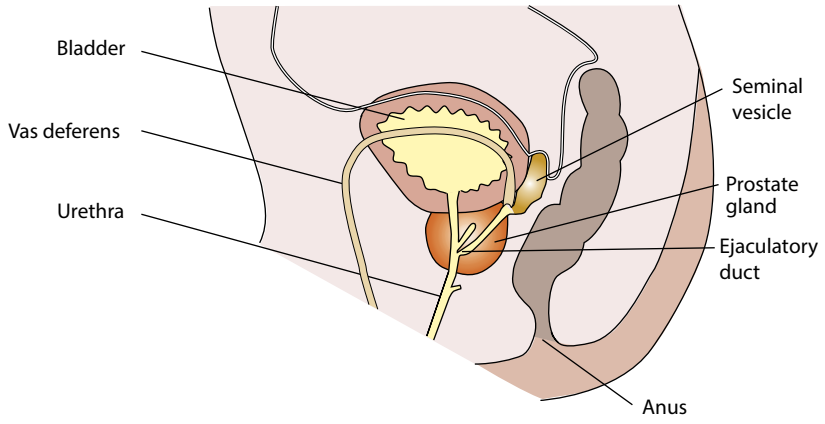


Figure 17.47 The prostate gland. ⏎

In the UK, more than 52 000 new cases of prostate cancer are diagnosed annually. It is the most common cancer in UK males and constitutes 14% of total cancers. It principally affects men over the age of 50 years, but its incidence is greatest in those aged 75–79 years. Approximately 12 000 UK deaths are attributable to prostate cancer. Its incidence increased over a number of years, which may have reflected the increasing age of the population and improved clinical testing for the condition, but in the UK it appears to be decreasing.

Signs, symptoms and diagnosis of prostate cancer

Many prostate cancers go undiagnosed because the tumour remains latent for long periods. While most prostate cancers are slow growing, a small proportion of tumours grow and metastasise more quickly. The risk of contracting prostate cancer are increased if:

- the man has mutated forms of *BRCA1* and *BRCA2* (*Section 17.4*)
- a female family member has had breast cancer, especially if diagnosed at an early age
- close relatives have suffered prostate cancer

Afro-Caribbean and African American men are at greater risk, while Asian men have the lowest risk of getting the disease. A diet high in animal fats and dairy products and low in fresh fruit and vegetables increases the risk. Consumption of the carotenoid lycopene (see *Figure 17.21*) reduces the risk of contracting prostate cancer. Lycopene has been shown to lower the amount of insulin-like growth factor 1 (*Chapter 7*), which may otherwise stimulate the growth of cancer cells.

The symptoms of prostate cancer are not usually present when the tumour is small. However, as the tumour grows, there is difficulty and pain when passing urine, coupled with a more frequent need to urinate, particularly at night. There may also be blood in the urine. If cancer cells have spread to the bone there may be pain in the back and pelvis. A patient presenting with symptoms of prostate cancer should be given a digital rectal examination (DRE), which involves inserting a gloved finger into the rectum to feel the prostate gland (*Figure 17.48*). An enlarged prostate that feels round and smooth is most likely a benign prostate tumour. In contrast, a malignant one makes the gland feel hard and lumpy. Blood samples are also taken for a prostate-specific antigen (PSA; *Figure 17.49*) test.

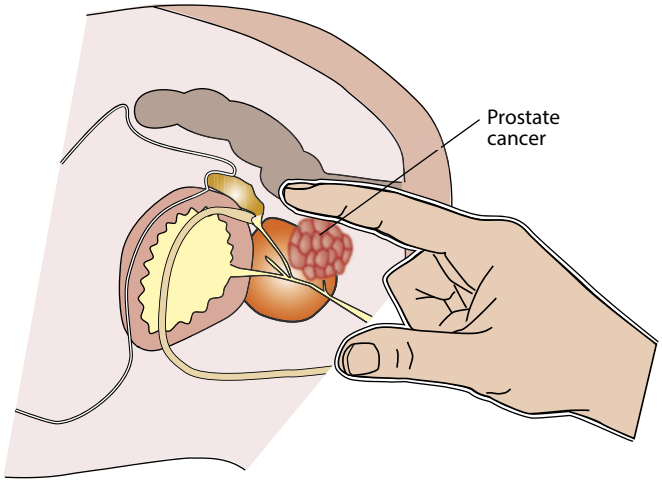


Figure 17.48 A digital rectal examination of a cancerous prostate gland. ⏎

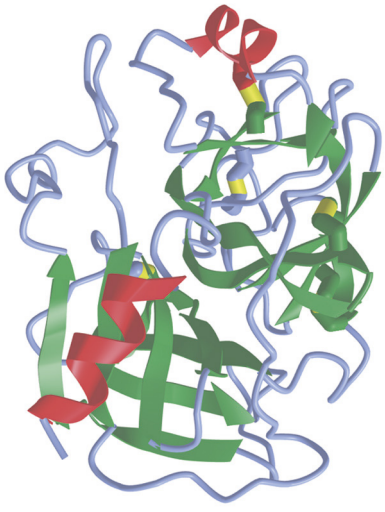


Figure 17.49 Molecular model of the prostate-specific antigen (PSA). PDB file 2ZCL. ⏎

The PSA test measures the concentration of PSA in blood. The normal concentrations of PSA are age related, and NICE values are listed in *Table 17.19*. Concentrations of PSA above the threshold for the age of the patient require referral for further tests because abnormal values are found in conditions other than cancer, for example in urinary infections. Following referral, additional tests to determine the extent of the problem include an isotopic scan of bones and an MRI or CT scan. Biopsies of the prostate are obtained by passing a needle through the rectum into the prostate, aided by ultrasound viewing.

Prostate cancers are commonly graded using a TNM system; however, a Gleason score describes its cell differentiation. A Gleason score is determined by assigning one Gleason grade to the most predominant cell pattern in the biopsy and a second grade to the second most dominant type. Grade 1 prostate cancer cells resemble normal prostate

Age (years)	Prostate-specific antigen threshold (µg/L)
Below 40	Use clinical judgement
40–49	>2.5
50–59	>3.5
60–69	>4.5
70–79	>6.5
Above 79	Use clinical judgement

Table 17.19 Age-specific prostate-specific thresholds to assist in diagnosing prostate cancer. ⏎

tissue, while those of grade 5 have mutated to such an extent that they have little resemblance to normal cells. The two grades are then added together to determine the Gleason score (*Table 17.20*). In theory, Gleason scores range from 2 to 10; however, pathologists almost never assign scores of 2–5 but give values of 6–10. A Gleason score of 6 is a low grade, 7 is an intermediate grade and 8–10 is a high grade prostate cancer.

Grade	Appearance of cells	Gleason score	Clinical viewpoint
1	Cells have a similar appearance to healthy prostate cells	6 (3+3)	Cancer is unlikely to grow or will grow very slowly
2	Majority of cells have a similar appearance to normal prostate cells	7 (3+4)	Cancer is likely to grow slowly
3	Cells look less like normal prostate cells	7 (4+3)	Cancer will probably grow at a moderate rate
4	Some cells look abnormal	8 (4+4)	Cancer may grow at a moderate rate or quickly
5	Cells look extremely abnormal	9–10 (4+5 / 5+4 / 5+5)	Cancer is likely to grow quickly

Table 17.20 The Gleason system of scoring prostate cancers. ⏎

Treatment for prostate cancer

The treatment for prostate cancer is determined by its TNM value and Gleason score, the age and health of the patient, and the concentration of PSA in the blood. Prostate cancers often grow slowly, and their treatment, which has considerable side effects, is often restricted to the more aggressive forms of the disease. The range of treatments for early prostate cancer include active surveillance, where the patient is regularly tested by DRE; measuring PSA blood concentrations; and MRI scanning to monitor the development of the prostate cancer. Other treatment options include radical prostatectomy (surgical removal of the prostate gland), radiation therapy, or combinations of any or all of them. Prostatectomy may result in impotence and incontinence. External beam radiation therapy and brachytherapy are also used as a first line of treatment. Conformational radiotherapy (CRT) allows the external radiation beam to be shaped to match that of the prostate, which lowers any side effects caused by irradiation of surrounding healthy tissues (see *Box 17.8*). Brachytherapy is delivered by the implantation of radioactive iodine 'seeds' or iridium wires in the prostate itself. If the cancer has spread to bone, radiation therapy using ^{89}Sr, which is preferentially taken up by bone tissue, is given by intravenous injection.

The growth of prostate cancer cells requires a supply of the biologically active form of testosterone (*Figure 7.13*) from the testes. Androgen deprivation therapy (ADT) is the first-line treatment for metastatic prostate cancer. It consists of orchiectomy (surgical removal of a testis) or hormonal therapy. Hormone therapy lowers the concentration of testosterone by using GnRH analogues, for example goserelin, leuprorelin and triptorelin. Goserelin is administered as a subcutaneous pellet, while the others are injected

subcutaneously or intramuscularly in liquid form. Castration-resistant prostate cancers, that is, those that do not respond to ADT, can be treated with antiandrogens, for example abiraterone, apalutamide, darolutamide and enzalutamide. Hormone therapy alone will not effect a cure, but it can relieve symptoms of the disease and slow or stop the growth and progression of the cancer. Also, not having surgery or radiotherapy means that the side effects of these treatments are avoided. However, the side effects of hormonal therapy can include sexual impotence, flushes and sweating, which may be reduced by intermittently stopping and starting the therapy.

Chemotherapy is rarely used to treat early prostate cancer. However, in the UK, the most common chemotherapeutics for treating advanced prostate cancer are cabazitaxel and docetaxel (*Figure 17.35*). Docetaxel may be used in combination with hormone therapy for men who have just been diagnosed with advanced prostate cancer or if hormone therapy has ceased to be effective. Cabazitaxel is used in patients whose advanced prostate cancer has stopped responding to hormone therapy and has failed to respond to docetaxel.

The prognosis for men with prostate cancer is generally good, since this type of cancer usually occurs in older men and grows slowly. In England, 90% of patients aged 15–64 years survive their disease for 10 or more years. Sadly, only about 25% of men aged 75–99 years live this long following diagnosis.

Leukaemias

Leukaemias are a heterogeneous group of haematological (*Chapter 13*) malignant neoplasms of bone marrow, lymph nodes, spleen and liver. More than 8000 leukaemias are diagnosed annually in the UK. They account for approximately 4800 yearly deaths. All leukaemias arise from the uncontrolled proliferation of a clone of a single type of blood cell progenitor and are characterised by the diffuse replacement of bone marrow with abnormally large numbers of immature leukocytes and a reduced number of circulating functional blood cells. Most leukaemias are characterised by excessive bleeding of the gums and nose, bruising, fatigue, breathlessness and increased susceptibility to infections. These are due to inadequate production of erythrocytes, platelets and lymphocytes owing to the abnormal proliferation of the clone of leukaemic cells in the bone marrow. Blood and bone marrow films may show immature blast cells (*Figure 17.50*).

The current classification of leukaemias uses a multiparameter method based on cell type and the extent of differentiation shown by the neoplastic cells, which, although comprehensive and useful, is rather complex for a text of this size. Hence, for simplicity, the historical, but still applicable, classification will be used here, that is: **acute** leukaemias, which have a sudden onset and rapid progression, and **chronic** types, which develop slowly over months and years, combined with *myelocytic* or *lymphocytic* based on whether the clone is lymphoid or myeloid in origin. The combination of fluorochrome dyes conjugated to antibodies and flow cytometry is a diagnostically useful way to distinguish between the two types of cells.

The main types of leukaemias encountered clinically are:

- acute lymphoblastic
- chronic lymphocytic
- acute myeloblastic
- chronic myelogenous

Other, rarer, forms of leukaemias do occur, including hairy cell leukaemia and acute promyelocytic leukaemia (*Margin Notes 17.17* and *17.18*, respectively).

Acute lymphoblastic leukaemia

Acute lymphoblastic leukaemias (ALLs) are classified into several different types, although two major forms are recognised depending on whether the leukaemic cells are T cells or B cells (B cell ALL and T cell ALL). However, the cause of both subtypes is unknown. Acute lymphoblastic leukaemias occur most frequently in children under the age of 15 years, peaking in 0 to 4 year olds, between the ages of 15 and 25 and in those over 75 years of age. Nearly 800 new cases occur annually in the UK.

The disease is characterised by the clonal proliferation of a lymphoid stem cell, leading to increased numbers of lymphoblasts in the blood. As with other leukaemias, patients may present with recurrent infections, bleeding of the gums and nose, bruising, anaemia, aching joints and bones, and a general feeling of malaise. Diagnosis is achieved by microscopic examination of blood smears showing increased numbers of immature lymphocytes combined with reduced numbers of erythrocytes and platelets. A bone marrow biopsy taken from the pelvis or sternum is essential for a definitive diagnosis. A lumbar puncture may be used to obtain a sample of cerebrospinal fluid, to detect

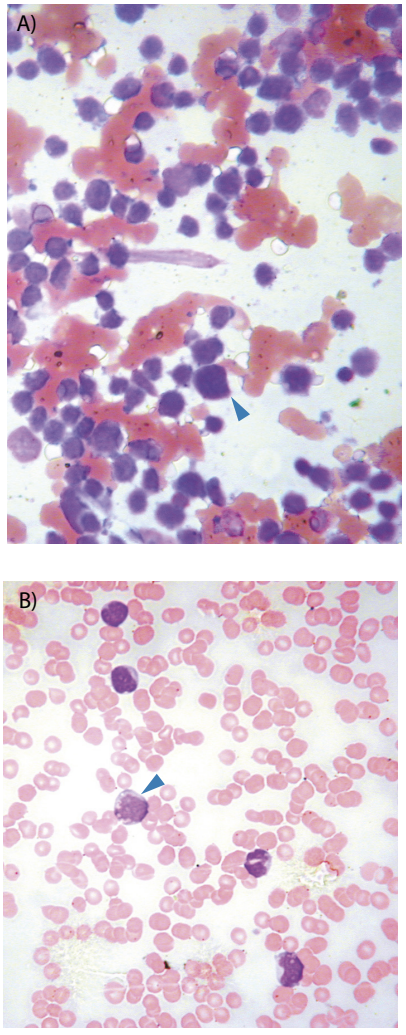


Figure 17.50 Blast cells, indicated by arrowheads, in (A) the bone marrow and (B) the blood of a patient with leukaemia.

the presence of leukaemic cells. The B cell and T cell forms can be distinguished using immunofluorescence testing of cell surface characteristics by flow cytometry (*Box 6.1* and *Chapter 4*). A proportion of ALL patients also have the Ph chromosome, which can be shown by karyotyping (see *Figures 17.12* and *17.13*).

The treatment for ALL involves chemotherapy. Steroids, such as dexamethasone (see *Figure 7.56*), may be combined with cytotoxic drugs. Chemotherapy may also require intrathecal injection of chemotherapeutic agents (*Section 17.7*) to destroy leukaemic cells in the cerebrospinal fluid. Radiation therapy to the brain and the testes in men may also be required to prevent recurrence. When patients are at high risk of relapse, an HSCT (*Chapter 6*) may be required. The survival rates for ALL depend on the age at which it is diagnosed. Children have a five-year survival rate of up to 80%; for adults it is only 20–35%.

Chronic lymphocytic leukaemia

Chronic lymphocytic leukaemia (CLL) is the most common form of leukaemia, affecting mostly those aged over 60–70 years. It is characterised by a clonal proliferation of lymphoid stem cells. However, the disease is slow growing, and many patients are asymptomatic in the early stages. Signs and symptoms include frequent viral and fungal infections, anaemia, bleeding and bruising, fatigue, weight loss and night sweats. The spleen and the lymph nodes in the neck, axillae and groin may be swollen due to the accumulation of abnormal cells. The cause of CLL is unknown, although in a minority of cases the disease may be familial.

A diagnosis of CLL is made by examination of blood and bone marrow. A lymph node biopsy may also be taken for histological examination. The UK uses the Binet staging system: CLL is described as stage A, B or C according to the degree of lymph node involvement. Patients with stage A, with little or no lymph node enlargement, may not need treatment. Stage B presents with three or more groups of enlarged lymph nodes, and stage C patients also have a low number of erythrocytes or platelets. The USA uses the Rai staging system, which divides the condition into five stages (0–IV).

Chronic lymphocytic leukaemia is treated by chemotherapy, although radiation therapy may be used to reduce any lymphadenopathy and splenomegaly. Patients with stages B and C may be brought into remission with chemotherapy. Treatment is usually successful in maintaining remission for several years. Monoclonal antibodies, such as the chimeric rituximab or the humanised alemtuzumab directed against the B lymphocyte surface antigens CD20 and CD52, respectively, are also available for therapy, although they can have pronounced side effects. The prognosis for CLL patients is variable, with survival times from diagnosis of 2–20 years. Low concentrations of platelets and haemoglobin at diagnosis are indicative of a poor prognosis.

Acute myeloid (myeloblastic) leukaemia

Acute myeloid leukaemia (AML) is a heterogeneous group of neoplastic disorders in which more than 20% of cells in the blood or bone marrow are immature myeloid cells (myeloblasts). It is a rare disease forming less than 1% of total cancer cases. In the UK, over 3000 adults are diagnosed with the condition annually, mostly over the age of 65 years, and peaking in the 80s. There is evidence that its incidence is increasing. Worldwide, the incidence of childhood AML is about seven per million.

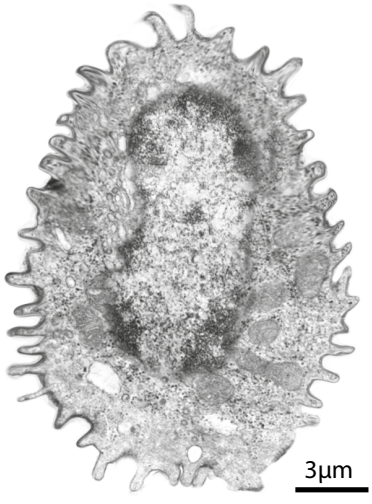


Figure 17.51 Electron micrograph of a single leukaemic cell from a patient with hairy cell leukaemia. Note the numerous extensions of the plasma membrane that give the condition its name. ⏎

Margin Note 17.17 ⏎
Hairy cell leukaemia

Hairy cell leukaemia (HCL) is a relatively rare form of chronic lymphocytic leukaemia. The incidence of HCL in the UK is not known due to its rarity. In 2019, it was estimated that nearly 6200 cases occurred in the USA, with over 1000 new cases expected yearly. It mainly presents in adults around 60 years old.
The leukaemic cell involved is an abnormal B lymphocyte characterised by outgrowths or projections, which give the disorder its name (*Figure 17.51*). Patients are prone to infections and present with symptoms of breathlessness, weakness and weight loss. There may also be splenomegaly.

A diagnosis is made by observing splenomegaly on examination, blood smears showing the presence of the distinctively shaped mononuclear HCL cells that can be up to double the size of a mature lymphocyte, and the presence of characteristic assembly of cell surface cluster of differentiation (CD) markers. Histological examination of a bone marrow specimen is also necessary.
This leukaemia is slow growing, and chemotherapy may only be needed if regular monitoring shows the number of abnormal cells is increasing and if the patient presents with symptoms.
First-line chemotherapy involves administering the purine analogues cladribine and pentostatin; both are equally effective in inducing and maintaining remission. Cladribine is increasingly used in combination with the chimeric monoclonal antibody rituximab, which binds to the cell surface antigen CD20 of B cells and initiates their death. However, interferon-α is preferentially used during pregnancy, and also may be the preferred initial therapy in patients with severe pancytopenia (a marked reduction in the formed elements of blood; *Chapter 13*).
Splenectomy may be considered in patients with symptomatic massive splenomegaly. Following treatment, bone marrow biopsies should be repeated to confirm complete remission, which is the complete absence of hairy cells in bone marrow and blood, resolution of the splenomegaly and recovery of peripheral blood counts. Hairy cell leukaemia is easily treatable (but not cured), with approximately 95% of patients surviving for five years or more after diagnosis.

Previously, AML was classified using features, such as cell morphology, but relatively recently, the WHO incorporated genetic abnormalities in a diagnostic algorithm that classified this diverse condition into six different categories:

1. AML with recurrent genetic abnormalities (for example translocations and inversions; *Chapter 15*)
2. AML with myelodysplasia-related changes (MRC)
3. therapy-related myeloid neoplasms (t-MN)
4. AML, not otherwise specified (NOS)
5. myeloid sarcoma
6. myeloid proliferations related to Down syndrome (*Chapter 15*)

A family history of AML is a positive risk factor. Also, numerous mutations, for example translocations and inversions, are associated with the condition. The cause of AML is unknown, although it is linked to exposure to industrial solvents, such as benzene, which is also found in tobacco smoke (see *Box 17.6* on tobacco abuse). Chemotherapy and therapeutic radiation treatment for other cancers are known risk factors.

Patients present with repeated infections, anaemia and bleeding from the gums and nose. They also bruise easily, due to inadequate production of leukocytes, erythrocytes and platelets, and suffer fatigue and lose weight. The condition requires rapid diagnosis and immediate treatment.

A diagnosis is achieved by the general clinical presentation and examining peripheral blood samples. Treatment is initially by chemotherapy, using a combination of cytarabine and daunorubicin (*Figure 17.52*). Cytarabine or cytosine arabinoside is a pyrimidine analogue that specifically targets rapidly dividing cells. Within cells it is converted to its triphosphate form and competes with cytidine for incorporation into replicating DNA. Once incorporated, replication is halted in the S phase of the cell cycle. Daunorubicin, or daunomycin, is an antibiotic. It acts by intercalating between the DNA base pairs, uncoiling the DNA double helix and inhibiting topoisomerase II, resulting in single- and double-stranded breaks. Hence, the biosyntheses of DNA and RNA are inhibited, stopping cancer cell growth. The humanised MAb (see *Box 4.5*) gemtuzumab, directed at CD33 proteins present on the surfaces of the leukaemic cells, is also used in combination with cytarabine, although, in the UK, it is not licensed for children and young people aged less than 15 years. Patients may require platelet transfusions before and during their treatment, and erythropoietin may be given to treat the anaemia. Complete remission is usually defined as the presence of normal blood cell counts with less than 5% blasts in the bone marrow. High-risk patients are recommended for an HSCT during their first remission.

Margin Note 17.18
Acute promyelocytic leukaemia

Acute promyelocytic leukaemia (APL; see *Box 15.3*) is a subtype of acute myeloid (myeloblastic) leukaemia. It is caused by a translocation between chromosomes 15 and 17 (t(15;17)(q24.1;21.2)), which produces the oncogene *PML-RARA* (*Figure 17.53*). The normal product of *PML*, the promyelocytic leukaemia protein, is a tumour suppressor. However, the oncoprotein PML-RARα prevents myeloblasts developing into myelocytes, with the result that promyelocytes accumulate in the bone marrow and blood. Symptoms of the condition include anaemia, a tendency to bruise easily and severe bleeding, and blood clots may also occur. Acute promyelocytic leukaemia is a serious clinical problem that needs treating as soon as possible.

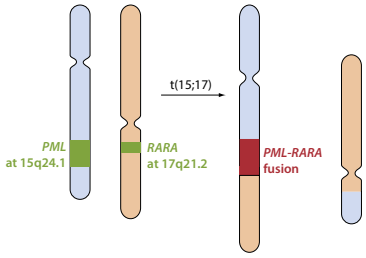


Figure 17.53 Schematic showing the translocation between chromosomes 15 and 17 that produces the fusion oncogene *PML-RARA* associated with acute promyelocytic leukaemia.

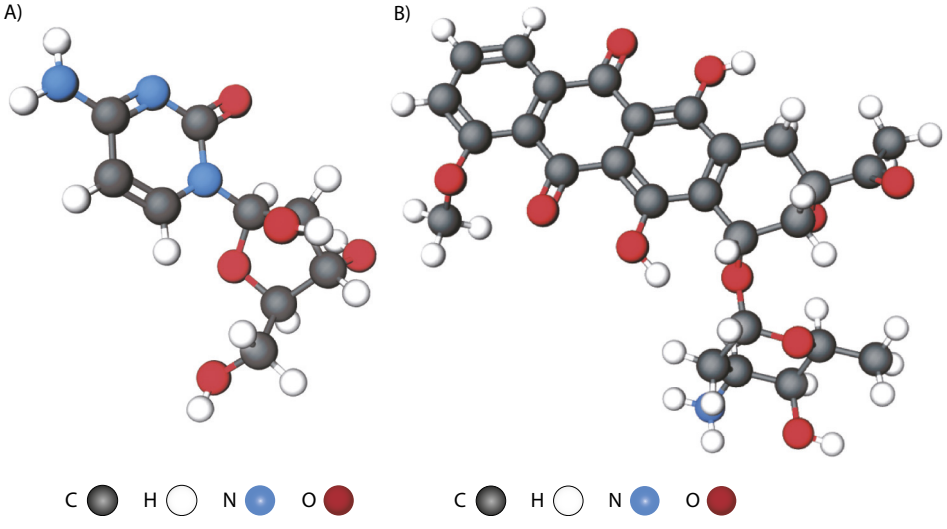


Figure 17.52 Computer drawn structures of (A) cytarabine and (B) daunorubicin.

The prognosis for AML depends on the patient age. Those younger than 40 years have about a 60% chance of surviving for five years following diagnosis. Unfortunately, the survival rate decreases rapidly with advancing years: it is more than 15% at 60–69, approximately 5% at 70–79 but only 1% at 80-plus years.

Chronic myelogenous leukaemia

Chronic myelogenous leukaemia (CML) is a rare disorder that accounts for less than 1% of all leukaemias in the UK, with approximately 800 new cases diagnosed annually. Patients are usually adults over 65 years old, with the incidence peaking at over 85 years.

The Ph chromosome, which contains the fusion oncogene *BCR-ABL* that encodes the oncoprotein BCR-ABL (see *Figures 17.12* and *17.13*), is present in 95% of CML patients. This malignant transformation in a myeloid stem cell, which would normally give rise to the polymorphonuclear leukocytes, monocytes, basophils, erythrocytes and megakaryocytes (*Chapters 4* and *12*), results in its clonal proliferation and the accumulation of immature and mature myeloid cells within the bone marrow and spleen.

Chronic myelogenous leukaemia usually follows a clinical pattern of three phases:

1. a benign chronic phase
2. an accelerated phase
3. a blastic or blast crisis phase

During the chronic phase, there are few signs and symptoms; indeed, many patients are asymptomatic. Patients may present with fatigue, splenomegaly and anorexia, and lose weight, with the condition only being revealed by a routine blood count during an investigation of these symptoms. Leukocytes in the blood increase in number (leukocytosis) from reference values of 10×10^9 to more than 500×10^9 cells/L. A blood smear also shows large numbers of neutrophils and immature myelocytes. Splenomegaly may be present on examination. The accelerated phase is indicated by an increased leukocytosis and the presence of myeloblasts in the blood. Progression to the blast crisis stage begins about 3–18 months following the beginning of the accelerated phase and may take three to five years from diagnosis. It is defined as more than 30% blasts in the bone marrow or peripheral blood, a deterioration in symptoms and the development of additional mutations. Blast crises are fatal after about three months.

A diagnosis of CML is initially based on general symptoms and microscopic examination of blood and bone marrow smears. This can be confirmed by molecular techniques, such as FISH (*Section 17.6*) or NGS methods (*Box 17.2*), and karyotyping to determine the presence of the Ph chromosome. The treatment for CML is chemotherapy with tyrosine kinase inhibitors (TKIs), such as bosutinib, dasatinib, imatinib and nilotinib (see *Table 17.14* and *Figure 17.14*). In most patients, TKIs control the disease and can effect a complete remission, that is, the ablation of the Ph chromosome and its product, the oncoprotein BCR-ABL. However, they do not *cure* the condition, which requires an HSCT (*Chapter 6*).

BOX 17.9 RETINOBLASTOMA ⏎

Retinoblastoma (Rb) is cancer of the retina. It is the most frequent malignant intraocular tumour in childhood. Worldwide, the incidence of Rb is approximately 1 in 17 000 live births, and about 8000 new cases are diagnosed annually. Some 40–50 cases are diagnosed in the UK each year. It usually affects children under five years old. Metastatic disease occurs in 10–15% of patients.

Two types of Rb occur: heritable forms, where tumours are often bilateral (present in both eyes); and nonheritable forms, with only one eye affected. Both types are associated with variants of the tumour suppressor gene *RB1* sited at 13q14.2 (*Section 17.4*). Inactivation of both alleles of *RB1* results in the production of a defective nuclear phosphoprotein, pRB, which in healthy individuals is a key regulator of the cell cycle. The biosynthesis of the oncoprotein form of pRB therefore impairs control of the cycle and produces an uncontrolled proliferation of neoplastic cells. In nonheritable Rb, both mutations in *RB1* occur in a single cell of the retina, which then grows to form the neoplasm. In heritable Rb, an early onset of the disease occurs if a germ line mutation is followed by a somatic mutation in *RB1*.

The first sign of Rb is often leukocoria, that is, a white pupil that does not reflect light. It is observable when a photograph of the child is taken using flash photography. The child may present with strabismus (a squint) or, if the tumour is large, they may have pain and bleeding from the eye. Screening of children with a familial history of Rb may detect a tumour before any symptoms appear.

A diagnosis of Rb requires ophthalmology tests to assess the size and position of the tumour, and whether it has metastasised. The Bruckner test simultaneously projects bright light through the pupils of both eyes. If Rb is present in an eye, then its pupil shines white (leukocoria; *Figure 17.54*). A fundus reflex examination is necessary to view and assess the position,

Figure 17.54 A positive Bruckner test showing leukocoria in the left eye. Courtesy of Isabella Greenwood and the Childhood Eye Cancer Trust (CHECT). ⏎

appearance and size of any Rb. Retinoblastomas appear as an undifferentiated white reflex or mass. Tests may include ultrasound and magnetic resonance imaging (MRI) scans and an optical coherence tomography (OCT) scan to form a detailed image of the retina. Blood bone marrow samples and an aqueous fluid sample (AFS) may be used to investigate tumour DNA. The extent of any metastases is investigated using a spinal fluid sample.

The commonly used staging system for Rb first divides them into intraocular or extraocular types. Intraocular Rbs can affect one or both eyes. This stage may be subdivided into five grades, depending on the size and position of the cancer and the extent of eye damage. The two most used staging systems are the International Intraocular Retinoblastoma Classification (IIRC; stages A–E) and the TNM system (stages cT1–cT4). Extraocular Rbs have metastasised from the eye(s) to the optic nerve, surrounding tissues or, indeed, more distal organs, for example the brain.

Treatment for Rb depends on the number, position and stage of the tumour(s). Naturally, the overriding aim is to cure the Rb but retain as much vision as possible. Local therapy includes cryotherapy and thermotherapy. In the former, a freezing appliance is attached to the outside of the eye in the region of the Rb to kill the neoplastic cells. Several treatments at monthly intervals may be necessary. Thermotherapy involves directing light from a laser through the pupil under general anaesthetic to heat and destroy the tumour. Again, several sessions of therapy may be required. Thermotherapy may be combined with chemotherapy or radiotherapy. Chemotherapy is usually used in combination with local therapy. The most used Rb directed drugs include carboplatin, cyclophosphamide, doxorubicin, etoposide, ifosfamide and vincristine (*Section 17.7*). Melphalan and topotecan, either alone or in combination, may be used in more advanced cases (*Figure 17.55*).

Radiation therapy is used to treat larger tumours or those that have resisted other therapies. It may involve attaching a small radioactive disc to the outside part of the eye that overlays the Rb. Irradiation may require up to four days to destroy the tumour, meaning the child must be hospitalised and isolated. Proton beam radiotherapy (*Section 17.7*) of Rb may be used after chemotherapy or when other treatments have been unsuccessful. Surgery is considered for advanced tumours and when vision has been lost. Enucleation is the complete removal of an eye together with a 10–15 mm length of the optic nerve. An artificial eye is then fitted.

Untreated Rb can be fatal within one to two years but, if adequately treated, the prognosis is excellent: over 95% of affected children are cured. However, the heritable form carries an increased risk of developing other tumours later in life. Families of, and children (when sufficiently mature) with, heritable Rb require sympathetic genetic counselling and regular monitoring into adulthood. Investigating any new symptoms and observing precautionary measures, such as avoiding tobacco abuse (*Box 17.6*), using protection from strong sunlight with appropriate sunglasses and a minimum of SPF30 ultraviolet blocking cream, restricting the use of alcohol, having a healthy diet (*Chapter 10*) and taking regular exercise, are essential.

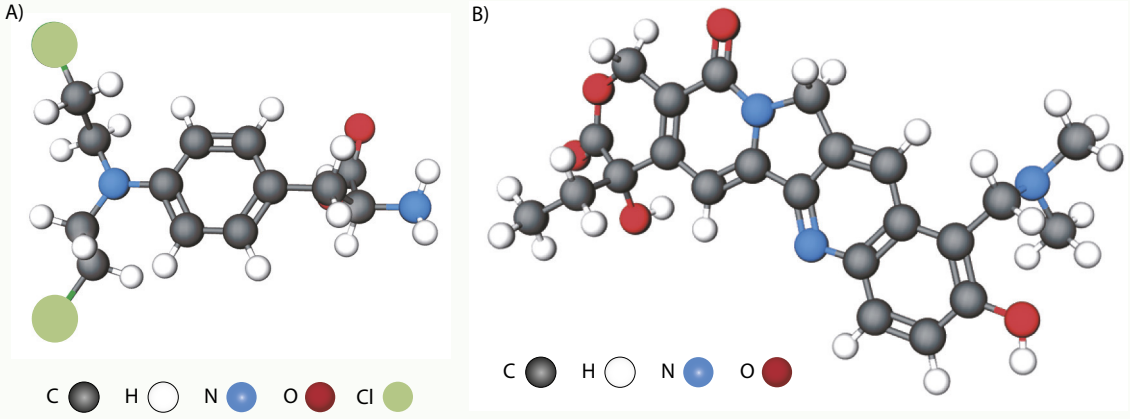


Figure 17.55 Computer drawn models of melphalan and topotecan.

17.9 CASE STUDIES

CASE STUDY 17.1

Chris, a 56-year-old man, recently noticed that he needed to urinate more frequently, especially during the night. Also, urination was painful. He visited his doctor, where a rectal examination showed his prostate gland was enlarged and smooth. A blood sample was taken for a PSA test. The results showed the level of PSA in his blood to be 5.3 μg/L.

Questions

(a) What is the likely diagnosis for Chris?
(b) What would be the recommended treatment?

CASE STUDY 17.2

Rebecca is a 55-year-old woman. About four weeks ago, she developed a bad cold and cough. While the cold cleared up, the cough has remained and has become more painful. During the last week, she was alarmed to find that she was coughing blood into her handkerchief, especially first thing in the morning. Rebecca gave up smoking about two years ago, but was previously a smoker of 20 cigarettes a day, starting from the age of 23 years. Her husband also smokes. Rebecca visited her physician who referred her to a consultant. The consultant saw Rebecca two weeks later and she had a full examination, including a chest X-ray.

Questions

(a) What clinical findings might the physician be expecting to see?
(b) What is the risk of Rebecca having lung cancer if she is a nonsmoker?

CASE STUDY 17.3

Giles is a 14 year old boy who, until recently, was a normal adolescent who loved sports, especially football. During the last few weeks, however, Giles has been feeling unwell. He has only played football twice in the last four weeks and, each time, he felt too tired during the first half to carry on for the whole match. Giles' mother noticed her son's tiredness and worried about the number of coughs and colds he suffered. However, when she noticed a number of large bruises on his legs and torso, she decided to take him to see his doctor, who took a sample of blood for a full blood and platelet count. The blood count indicated a leukocytosis, with numerous blast cells present. Platelet numbers were reduced and there were indications of anaemia.

Questions

(a) What disorders might be suspected?
(b) What further tests would confirm or refute your suspicions?

CASE STUDY 17.4

Eric, a 17-month-old toddler, had a swollen right eye. The eye movements were uncontrolled, and Eric had a strabismus (squint) in the eye. The boy was taken to see an ophthalmologist. A vision test revealed the right eye had reduced visual acuity and tracking ability. A Bruckner test showed a red reflex in his left eye but leukocoria in the right. A fundus reflex examination of the left eye gave a normal result. The right eye had a clear vitreous humour and a pinkish-white, spherical mass in the retina. An MRI scan of the right eye detected an intraocular tumour presenting with calcification. All extraocular structures were normal. A systemic examination of Eric showed no abnormalities. There was no family history of ocular problems.

Questions

(a) What is the most plausible diagnosis of Eric's condition?
(b) How could the underlying cause of this diagnosis be investigated?
(c) Given a confirmed diagnosis, how should Eric be treated?

CASE STUDY 17.5

Helen, a 47 year old, found a lump in her right breast while having a shower. She had no known family history of breast or ovarian cancer. An ultrasound examination and a mammography revealed a roughly circular mass of approximately 16–18 mm diameter with spiculated margins in her right breast. The initial diagnosis was of a poorly differentiated invasive ductal carcinoma. No definitive lymphovascular invasion or carcinoma *in situ* were identified. Immunohistochemical staining of biopsy material showed the tumour cells were negative for oestrogen and progesterone receptors, while the result for *HER2* overexpression was equivocal. A FISH test was requested to determine Helen's *HER2* status. *Figure 17.56* shows the result of the FISH investigation.

Questions

(a) Assign a staging to Helen's cancer.
(b) How does the status of Helen's oestrogen and progesterone receptors and *HER2* amplification tests inform her treatment?
(c) Suggest how Helen should be treated.
(d) What is the most likely prognosis for Helen?

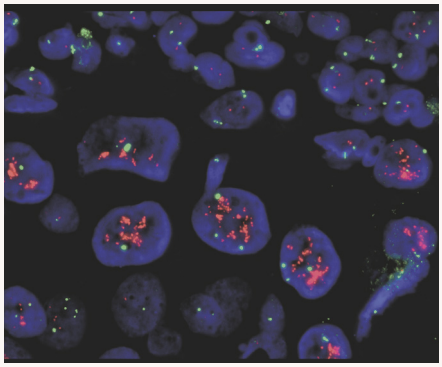


Figure 17.56 The use of fluorescence *in situ* hybridisation (FISH) in diagnosing *HER2* positive breast cancer. This test uses two probes to detect *HER2* amplification. The probe for *HER2* gives an orange fluorescence, while the internal control probe for the centromere of chromosome 17 (*CEP17*) has a green fluorescent signal. Courtesy of Dr Nick Telford and Clare Hodgson, Oncology Cytogenetics, The Christie Pathology Partnership, The Christie NHS Foundation Trust, Manchester, UK. ⏎

17.10 SUMMARY

Cancer is a group of diseases in which cells escape from the usual regulatory factors controlling cell division and proliferate to form tumours. Cells from tumours may break free and move through the blood and lymph to establish tumours at distant sites. Cancer is ultimately caused by successive mutations in the DNA and epigenetic changes that confer a selective advantage to the cancer cells. Examples of genes that, when mutated, may give rise to cancers include cellular protooncogenes and tumour suppressor genes. Mutations may be caused by chemical carcinogens or radiation, by oncogenic viruses or by intrinsic failures of DNA repair mechanisms. Some gene mutations that predispose to cancer are inherited. Examples include mutated forms of the *BRCA1* and *BRCA2* genes, which predispose women to cancers of the breast and/or ovaries.

Diagnosis of cancer depends on the location and symptoms produced by the tumour. A whole battery of diagnostic procedures may be used, including tumour imaging, histological examination of tumour biopsies, liquid biopsies, FISH and immunoassays for tumour associated markers. Treatments for cancers include chemotherapy, radiation therapy, hormone therapy, immunotherapy, and targeted and gene therapies.

QUESTIONS

1. Which of the following is the odd ONE out?
(a) Epstein–Barr virus.
(b) HIV.
(c) Kaposi virus.
(d) HPV.
(e) Human T cell leukaemia virus.

2. Which of the following statements is incorrect?
(a) A benign tumour does not metastasise.
(b) Oncogenes may be found in certain viruses.
(c) Tumour suppressor genes control cell division.
(d) Cancer may be treated with cytotoxic drugs.
(e) Leukaemic cells originate in the blood.

3. Which of the following are the four most common tumour sites in adults in the UK?
(a) Lung, breast, colorectum and bone marrow.
(b) Breast, lung, colorectum and prostate.
(c) Lung, breast, stomach and prostate.
(d) Brain, lung, breast and bone marrow.
(e) Stomach, lung, skin and breast.

4. Arrange the following two lists into their most appropriate pairings.

Carcinoma	*BCR-ABL*
Metastasis	*TP53*
Oncogene	Myeloblastic leukaemias
Tumour suppressor gene	Hodgkin lymphoma
Breast cancer	Xeroderma pigmentosum
Li–Fraumeni syndrome	Secondary tumours
Tyrosine kinase	*EMSY*
FAB	Epithelia
Ultraviolet light	*E2F3*
Epstein–Barr virus	*CHK2*

5. Indicate whether the following statements are TRUE or FALSE.
(a) Cachexia is defined as an unexplained loss of 2 kg of body weight.
(b) A viral equivalent of an abnormal human gene that is associated with cancer is denoted by the prefix *v-*.
(c) Cigarette smoke contains benzo[a]pyrene, which is a direct acting carcinogen.
(d) The use of a cryostat allows for a rapid histological examination of biopsy material from a suspected cancer.
(e) Cells undergoing mitosis are at their most resistant to radiation damage.
(f) Methotrexate is an antagonist to folic acid synthesis.
(g) Cyclophosphamide and 6-TA are used in the treatment of breast cancer.
(h) Vincristine, obtainable from *Colchicum officinale*, is used to treat leukaemias.
(i) Coley's toxin contains killed mycobacteria that stimulate the immune system.
(j) Hairy cell leukaemia is the name of a rare form of ALL.

6. Which of the following statement(s) is/are FALSE?
   (a) Stress signals can initiate the intrinsic apoptotic pathway through the activities of Bcl-2 proteins.
   (b) Initiator caspases include caspases-8, -7 and -9.
   (c) Fas, Fas ligands and the Fas-associated death domain protein all possess death domains.
   (d) The condensation of DNA, cell shrinkage and blebs of the plasma membrane are characteristics of an apoptotic cell.
   (e) Apoptosis is an unnatural form of cell death.

7. Carcinomas constitute 80–90% of all human cancers. Suggest why carcinomas are the most common types of tumours.

8. Which ONE of the following is not a hallmark of most cancers?
   (a) Sensitivity to growth inhibitory signals.
   (b) Growth signal autonomy.
   (c) Invasion and metastasis.
   (d) Sustained angiogenesis.
   (e) Evasion of immune surveillance.

9. Which of the following types of cancer is most likely to be described using a TNM staging system?
   (a) Hairy cell leukaemia.
   (b) Chronic myelogenous leukaemia.
   (c) Multiple myeloma.
   (d) Non-small cell lung cancer.
   (e) Chronic lymphocytic leukaemia.

10. Which of the following imaging test(s) would be performed to determine if prostate cancer has spread to the bones, lungs or brain?
    (a) Ultrasound.
    (b) X-ray.
    (c) MRI.
    (d) CT scan.
    (e) PET scan.

FURTHER READING

Abaza A *et al*. (2022) A systematic review of apoptosis in correlation with cancer: should apoptosis be the ultimate target for cancer treatment? *Cureus* **14**:e28496. doi: 10.7759/cureus.28496.

Abdel-Wahab M *et al*. (2021) Global radiotherapy: current status and future directions – white paper. *JCO Glob Oncol* **7**:827–842. doi: 10.1200/GO.21.00029.

Afzal S *et al*. (2022) Breast cancer: discovery of novel diagnostic biomarkers, drug resistance, and therapeutic implications. *Front Mol Biosci* **9**:783450. doi: 10.3389/fmolb.2022.783450.

Arber DA *et al*. (2022) International consensus classification of myeloid neoplasms and acute leukemias: integrating morphologic, clinical, and genomic data. *Blood* **140**:1200–1228.

Basudan AM (2022) The role of immune checkpoint inhibitors in cancer therapy. *Clin Pract* **13**:22–40.

Bruni L *et al*. (2023) Global and regional estimates of genital human papillomavirus prevalence among men: a systematic review and meta-analysis. *Lancet Glob Health* **11**:e1345–e1362.

Butterfield LH and Najjar YG (2024) Immunotherapy combination approaches: mechanisms, biomarkers and clinical observations. *Nat Rev Immunol* **24**:399–416.

Carneiro BA and El-Deiry WS (2020) Targeting apoptosis in cancer therapy. *Nat Rev Clin Oncol* **17**:395–417.

Chen X *et al*. (2022) Mutant p53 in cancer: from molecular mechanism to therapeutic modulation. *Cell Death Dis* **13**:974. doi: 10.1038/s41419-022-05408-1.

Cheng Y *et al*. (2019) Targeting epigenetic regulators for cancer therapy: mechanisms and advances in clinical trials. *Signal Transduct Target Ther* **4**:62. doi: 10.1038/s41392-019-0095-0.

Choi HY and Chang JE (2023) Targeted therapy for cancers: from ongoing clinical trials to FDA-approved drugs. *Int J Mol Sci* **24**:13618. doi: 10.3390/ijms241713618.

Connal S *et al*. (2023) Liquid biopsies: the future of cancer early detection. *J Transl Med* **21**:118. doi: 10.1186/s12967-023-03960-8.

Cruz-Gálvez CC *et al*. (2022) Retinoblastoma: review and new insights. *Front Oncol* **12**:963780. doi: 10.3389/fonc.2022.963780.

Debela DT *et al*. (2021) New approaches and procedures for cancer treatment: current perspectives. *Sage Open Med* **9**:20503121211034366. doi: 10.1177/20503121211034366.

Doll SR (2000) Smoking and lung cancer. *Am J Respir Am J Crit Care Med* **162**:4–6.

Farha MW and Salami SS (2022) Biomarkers for prostate cancer detection and risk stratification. *Ther Adv Urol* **14**:17562872221103988. doi: 10.1177/17562872221103988.

Farina E *et al*. (2022) An overview of artificial intelligence in oncology. *Future OA* **8**:FSO787. doi: 10.2144/fsoa-2021-0074.

Filella X *et al*. (2023) Clinical usefulness of circulating tumor markers. *Clin Chem Lab Med* **61**:895–905.

Finger PT and Tomar AS (2022) Retinoblastoma outcomes: a global perspective. *Lancet Glob Health* **10**:e307–e308.

Flaitz CM and Hicks MJ (1998) Molecular piracy: the viral link to carcinogenesis. *Oral Oncol* **34**:448–453.

Gamble P *et al*. (2021) Determining breast cancer biomarker status and associated morphological features using deep learning. *Commun Med (Lond)* **1**:14. doi: 10.1038/s43856-021-00013-3.
Gibbs SN *et al*. (2023) Comprehensive review on the clinical impact of next-generation sequencing tests for the management of advanced cancer. *JCO Precis Onco* **7**:e2200715 doi: 10.1200/PO.22.00.
Guallar-Garrido and Julián E (2020) Bacillus Calmette–Guérin (BCG) therapy for bladder cancer: an update. *ImmunoTargets Ther* **9**:1–11.
Harbeck N *et al*. (2019). Breast cancer. *Nat Rev Dis Primers* **5**:66. doi: 10.1038/s41572-019-0111-2.
Hoffbrand AV *et al*. (2019) *Essential Hematology*, 8th edition. Wiley Blackwell Science.
ICGC/TCGA Pan-Cancer Analysis of Whole Genomes Consortium (2020). Pan-cancer analysis of whole genomes. *Nature* **578**:82–93.
Krump NA and You J (2018) Molecular mechanisms of viral oncogenesis in humans. *Nat Rev Microbiol* **16**:684–698.
Kumar N and Sethi G (2023) Telomerase and hallmarks of cancer: an intricate interplay governing cancer cell evolution. *Cancer Lett* **578**:216459. doi: 10.1016/j.canlet.2023.216459.
Lawrence R *et al*. (2023) Circulating tumour cells for early detection of clinically relevant cancer. *Nat Rev Clin Oncol* **20**:487–500.
Leiter A *et al*. (2023) The global burden of lung cancer: current status and future trends. *Nat Rev Clin Oncol* **20**:624–639.
Leong MML and Lung ML (2021) The impact of Epstein–Barr virus infection on epigenetic regulation of host cell gene expression in epithelial and lymphocytic malignancies. *Front Oncol* **11:**629780. doi: 10.3389/fonc.2021.629780.
Liao J *et al*. (2022) Artificial intelligence assists precision medicine in cancer treatment. *Front Oncol* **12**:998222. doi: 10.3389/fonc.2022.998222.
Liesner RJ and Goldstone AH (1997) ABC of clinical haematology. The acute leukaemias. *BMJ* **314**:733–736.
Lipsick J (2021) A history of cancer research: tumor viruses. *Cold Spring Harb Perspect Biol* **13**:a035774. doi: 10.1101/cshperspect.a035774.
Lorentzen CL *et al*. (2022) Clinical advances and ongoing trials on mRNA vaccines for cancer treatment. *Lancet Oncol* **23**:e450–e458.
MacLennan SA and Marra MA (2023) Oncogenic viruses and the epigenome: how viruses hijack epigenetic mechanisms to drive cancer. *Int J Mol Sci* **24**:9543. doi: 10.3390/ijms24119543.
Mui UN *et al*. (2017) Viral oncology: molecular biology and pathogenesis. *J Clin Med* **6**:111. doi: 10.3390/jcm6120111.
Nelson CW and Mirabello L (2023) Human papillomavirus genomics: understanding carcinogenicity. *Tumour Virus Res* **15**:200258. doi: 10.1016/j.tvr.2023.200258.
Nicholson BD *et al*. (2020) The association between unexpected weight loss and cancer diagnosis in primary care: a matched cohort analysis of 65,000 presentations. *Br J Cancer* **122**:1848–1856.
Reinhardt D *et al*. (2022) Pediatric acute myeloid leukemia – past, present, and future. *J Clin Med* **11**:504.
Sarhadi VK and Armengol G (2022) Molecular biomarkers in cancer. *Biomolecules* **12**:1021. doi: 10.3390/biom12081021.
Satam H *et al*. (2023) Next-generation sequencing technology: current trends and advancements. *Biology* **12**:997. doi: 10.3390/biology12070997.
Sharma P *et al*. (2023) Immune checkpoint therapy – current perspectives and future directions. *Cell* **186**:1652–1669.
Shorokhova MT *et al*. (2021) Chromothripsis – explosion in genetic science. *Cells* **10**:1102.
Teichgraeber DC *et al*. (2021) Breast cancer staging: updates in the *AJCC Cancer Staging Manual*, 8th Edition, and current challenges for radiologists, from the *AJR* special series on cancer staging. *AJR Am J Roentgenol* **217**:278–290.
Troussard X (2022) Hairy cell leukemia 2022: update on diagnosis, risk-stratification, and treatment. *Am J Hematol* **97**:226–236.
Waarts MR *et al*. (2022) Targeting mutations in cancer. *J Clin Invest* **132**:e154943. doi: 10.1172/JCI154943.
Weinberg RA (2023) *The Biology of Cancer*, Third International Student Edition. WW Norton & Company.
Weth FR *et al*. (2024) Unlocking hidden potential: advancements, approaches, and obstacles in repurposing drugs for cancer therapy. *Br J Cancer* **130**:703–715.
Zapatka M *et al*. (2020) The landscape of viral associations in human cancers. *Nat Genet* **52**:320–330.

Useful websites

www.cancer.gov
www.cancer.org
www.cancerresearchuk.org
www.cdc.gov/cancer
https://portal.gdc.cancer.gov
www.nice.org.uk
www.who.int/health-topics/cancer#tab=tab_1
https://emedicine.medscape.com

Taylor & Francis
Taylor & Francis Group
http://taylorandfrancis.com

Ageing and disease

18

OBJECTIVES

After studying this chapter you should be able to:

- define several age-related terms
- describe some of the effects of ageing on cells, tissues, organs and systems
- list the possible causes of the effects of ageing
- describe several age-related clinical conditions
- discuss how some of the clinical and social problems of ageing may be managed

18.1 INTRODUCTION

Ageing is difficult to define, but is perhaps most widely understood to be a decrease in the ability to survive. As people age (*Figure 18.1*), they are less able to perform strenuous physical activities that were relatively easy when they were younger. With ageing comes a decline in the function of most organs in the body, making elderly people more susceptible to a variety of diseases. Indeed, most major diseases of the developed world, such as coronary heart disease (*Chapter 14*), cancer (*Chapter 17*) and diabetes type 2 (*Chapter 7*) are age-related. With ageing comes **senescence**, that is, a decline in functions of almost all parts of the body and at all levels of organisation, from cells to organ systems. Senescence changes may be responsible for some diseases of old age or may increase susceptibility to certain diseases. Thus, elderly people make up a large proportion of patients in hospitals and, in most countries, a high proportion of the health budget is devoted to the care and treatment of the elderly.

The mean length of life of individuals in a population is their **life expectancy**. Mean life expectancy has increased over the past 100 years, for example in the UK from 46 in 1893 to 82 years in 2023. For unknown reasons, there is a difference in the life expectancy of males and females, with women living longer than men: in 2023 in the UK, life expectancies for women and men were 83.1 and 79.2 years, respectively. Over the years, increases in life expectancy have been due to improvements in public healthcare, housing (*Chapter 2*) and nutrition (*Chapter 10*) in addition to developments in clinical care. The increased life expectancy has resulted in an enormous expansion of the elderly population in all industrialised countries, although this is less marked in developing countries. However, despite increases in life expectancy, the **human lifespan**, that is, the maximum age that can be attained, has not lengthened. Humans can live up to the age of about 120 years, but such longevity is exceptional. Humans live longer than other mammals: an elephant, for example, has a lifespan of 70 years, while that of a mouse is a mere three years. While the lifespan of species is inherent, humans are able, to a certain extent, to increase their life expectancy by controlling their environment.

Figure 18.1 The characteristic appearances of an elderly couple and their young grandchild.

Margin Note 18.1
Studying ageing

Most academic studies on ageing and attempting to prolong life have used small mammals, such as rats and mice. Consequently, more is known of such processes in these species than in humans.

18.2 AGEING OF CELLS, TISSUES, ORGANS AND SYSTEMS

The deleterious effects of the ageing processes are numerous and diverse. They affect all aspects of the body: cells, tissues, organs and systems.

Cells

Cellular functions decline in efficiency with advancing age. For example, the abilities of mitochondria to survive a hypoxic insult and perform oxidative phosphorylation, the

DOI: 10.1201/9781003296836-18

biosyntheses of structural, enzyme and receptor proteins, and the abilities of cells to take up nutrients and repair chromosomal damage all decline with age. Aged cells have irregular and abnormally shaped organelles, particularly their nuclei, Golgi apparatus and endoplasmic reticulum (*Chapter 16*), and also accumulate excessive waste products.

Tissues

All tissues are affected during ageing. For example, muscle mass is subject to **atrophy** with a reduction in the sizes of muscle groups due to the loss of individual muscle fibres. This results in a decreased capacity for work. Other factors, such as cardiovascular, respiratory and joint functions, also influence muscle strength. If the elderly are disabled by disease, for example arthritis (*Chapter 5* and *Section 18.4*), mobility may also be restricted, and muscles will undergo atrophy unless specific exercises are undertaken.

Organs

Age-related changes to organs include a decrease in the size and activities of several major organs. There is a decrease, for example, in both the size and the elasticity of the lungs, resulting in a reduced gas exchange capacity (*Chapter 5*). In general, the function of the lungs is still sufficient for most activities, although the capacity for strenuous activity is restricted due to a decline in cardiovascular function (*Chapter 14*). The weight and volume of kidneys may decrease 20–30% with age as nephrons are lost and replaced by scar tissue, which decreases the rate of ultrafiltration and therefore excretory capacity (*Chapter 8*). Hence, older people are more at risk of developing renal disease. The liver also shrinks in size due to loss of cells. There is a concomitant decline in some liver functions, especially in the metabolism and detoxification of drugs and xenobiotics (*Chapter 12*). This is clinically significant because it means that many medications are metabolised and cleared from the body more slowly, a fact that must be considered when prescribing drugs for the elderly.

Loss of sense organ functions occurs during ageing. A decline in the ability of the lens to change shape makes focusing on near objects more difficult, and the wearing of spectacles becomes a necessity. Lens crystallins (*Figure 18.2*) become more cross-linked and browner in colour, absorbing and scattering the light entering the eye; thus, less light reaches the retina. A deterioration in vision due to the formation of cataracts in the lens increases with advancing years (*Section 18.4*). With age, there is also a general physiological deterioration of the auditory system, particularly in the perception of higher frequencies.

The brain loses weight with age, shrinking from a typical mass of 1.4 kg at 20 to about 1.3 kg at 60 years of age. The loss is due to changes in composition that include an enlargement of the ventricles and a widening of the surface channels. Neurons are also lost, and amyloid protein may be deposited (see *Box 18.2*). An accumulation of the pigment lipofuscin (see *Figure 18.9*) also occurs in certain neurons. These changes are believed to contribute to a lengthening in reaction times, a decline in problem-solving and learning abilities, and an impairment of memory.

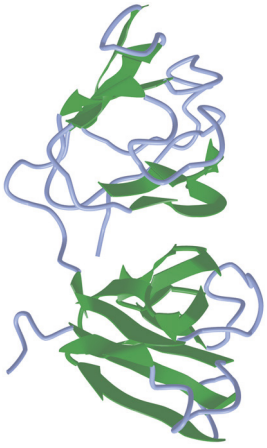


Figure 18.2 Molecular model of a γ crystallin. PDB file 2M3T. ⏎

Systems

Changes to the skin are among the most easily recognised effects of ageing (*Figure 18.3*). Indeed, many people use the appearance of skin and hair to assess the age of an individual. These changes include wrinkling, changes in skin pigmentation, and greying and loss of hair. Skin wrinkling is caused by changes to collagen, which has increased cross-links and reduced elasticity. The follicles of the skin produce hairs lacking pigment-forming melanocytes and so appear grey. There is a large variation in hair loss, but this is not surprising, given the many genetic and hormonal influences involved. Skin wounds heal more slowly in older individuals. Studies have compared healing of ischaemic (reduced blood flow) and fully vascularised wounds in young and old rats. The fully vascularised wounds healed equally well in both populations. However, healing of ischaemic wounds took significantly longer in the older animals. It may be that impaired wound healing in older people is related to conditions such as atherosclerosis (*Chapter 14*), which contributes to ischaemia of wounded tissues.

Endocrine function declines with age, both in reduced hormone production and in the number of hormone receptors on target cells. This results in, for example, an age-related decline in the functions of the reproductive organs, alterations in thyroid hormone status and an increased risk of developing diabetes (*Chapter 7*).

Immune functions also decline with age. The thymus atrophies and there is a progressive decline in the function of T lymphocytes (*Chapter 4*). There is a concomitant decrease in the production of antibodies, possibly due to loss of regulatory T cell

Figure 18.3 The hand of an elderly person. Note the wrinkled skin and brown ('liver') spots. ⏎

BOX 18.1 SUNLIGHT AND SKIN

Type I collagen is the major collagen (*Figure 18.4*) in the dermis of the skin. Its destruction, along with damage to other structural components of the skin, occurs over decades and is thought to underlie the characteristic alterations in the appearance of aged skin (*Figure 18.3*). Ultraviolet (UV) light from the sun is a major factor contributing to the premature ageing of skin. This effect, called photoageing, presents with fine and coarse wrinkling, blotchy pigmentation and a rough skin texture often described as 'leathery'. These sun-induced changes are additional to intrinsic ageing processes. For this reason, people who spend long periods in the open air, either through occupation or choice, are strongly advised to use protective UV light-blocking creams. The healthiest skin is found on areas usually well covered, such as that of the genitalia and buttocks.

The mechanisms underlying collagen degradation in photoageing are not fully understood but are at least in part due to the activity of matrix metalloproteinases (MMPs; see *Chapter 17*) released from keratinocytes and fibroblasts. The levels of these collagen-degrading proteases in skin increase as a function of age and are also transiently increased in response to the action of UV light. In addition, the biosynthesis of new collagen decreases as ageing progresses. All these effects mean there is increased damage to and loss of function of the collagen in the dermis. The difference between young and old skin is, in part, a reflection of an intrinsic reduction in the capacity of old fibroblasts to biosynthesise collagen.

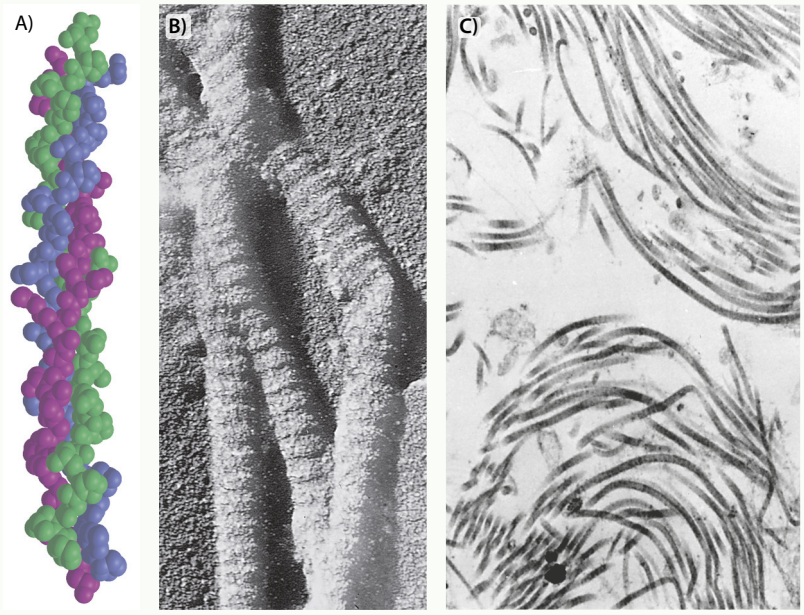


Figure 18.4 (A) Molecular model of a portion of a collagen molecule. PDB file 1BKV. (B) Electron micrograph of collagen. (C) Collagen fibres in a sample of tissue. ⏎

Overexposure of the skin to UV light, from the sun and from tanning beds, is also associated with an increase in skin cancer, especially malignant melanoma, a particularly aggressive form of cancer (*Chapter 17*) that can also occur in the pigmented retina of the eye.

function, together with an increase in the development of autoimmune reactions. This increases susceptibility to infections in the elderly, coupled with autoimmune injury to cells and tissues.

Numerous well-documented defects may occur in the cardiovascular system as it ages. Connective tissues, which are essential components of blood vessel walls, lose elasticity and this increases the rigidity of the vessels. Blood vessels are also prone to calcification and hardening of the arteries, leading to atherosclerosis (*Chapter 14*). The narrowing of the lumen of blood vessels by arteriosclerosis leads to an increase in blood pressure in the elderly. With age, the cardiac muscle also becomes less efficient and the heart enlarges due to accumulating fibrotic tissue, leading to a decline in cardiac output. A consequence of these changes is a reduced delivery of blood to peripheral tissues and to the heart itself.

18.3 CAUSES OF AGEING

Several theories have been proposed to explain cellular ageing. These theories can be divided into two broad groups, those that are:

1. based on 'wear and tear'
2. genome-based

'Wear and tear' theories

Wear and tear theories suggest that ageing processes in cells are due to a continual exposure to harmful agents from both inside and outside the cell throughout life. These agents include:

- free radicals
- glycated proteins
- waste products and the products of erroneous biosyntheses
- the error-catastrophe theory with the accumulation of abnormal proteins

Free radicals

Free radicals are molecules that possess an unpaired electron, which makes them extremely chemically reactive. Most free radicals in the body are derivatives of dioxygen and are often referred to as reactive oxygen species (ROS). Indeed, the most common free radical in the human body is the superoxide radical ($O_2^{\bullet-}$). Superoxide radicals can form hydroxyl radicals ($OH^{\bullet}$), which are the most reactive ROS.

Sources of free radicals can be endogenous, that is, formed within the body, or exogenous, when they are supplied from outside the body. Free radicals are produced by the electron transport chain of the inner mitochondria membrane (*Chapter 16*) where dioxygen normally forms water by accepting electrons and protons. However, a relatively small number of dioxygen molecules accept only a single electron and form superoxide radicals (*Chapter 12*). Free radicals are also formed by the interactions of ionising radiation and water in the body, which can give rise to hydroxyl radicals, and in ischaemia when an inadequate blood flow is followed by reperfusion (restoration of blood flow) as in a myocardial infarction (*Chapter 14*). The one electron reductive catabolism of xenobiotics, for example paraquat, carbon tetrachloride, quinone and paracetamol, which occurs in the liver and involves dioxygen, can also produce free radicals, such as superoxide and subsequently hydroxyl radicals (*Chapter 12*).

It is estimated that each body cell is attacked by approximately 10 000 free radicals daily. They are thought to react with biomolecules forming new free radicals and thus initiate damaging chain reactions (see, for example, *Figure 12.25*). For example, free radicals can react with lipids, forming lipid peroxides that can damage cell membranes, making them 'leaky' and causing their breakdown. Free radicals may also fragment proteins and when generated near DNA can react with its bases, producing mutations.

The reactivity of free radicals can be mitigated if two radicals meet and form a covalent bond or by the donation of electrons to or the removal of electrons from other molecules. Antioxidants are molecules that can protect against oxidation by ROS. They are found both within and outside cells. Antioxidants can act by:

- decreasing or removing dioxygen
- removing catalytic metal ions
- scavenging free radicals or their precursors
- breaking chain reactions initiated by free radicals

In general, antioxidants can be divided into two groups: extracellular and intracellular types. Extracellular antioxidants tend to be dietary in origin and include free radical scavengers, such as vitamins C and E (see *Figures 10.18* and *10.21*), which being lipid soluble may protect biological membranes against damage (*Chapter 10*). There is evidence from studies using rats, mice and some nonmammalian species that such dietary antioxidants may delay the ageing process and increase life expectancy, but it is not known whether they act solely by reducing free radical damage. Intracellular antioxidants include reduced glutathione (GSH), a tripeptide present in most cells, which contains a thiol (-SH) group that is readily oxidised, giving rise to oxidised glutathione (GSSG) (*Figure 18.5*). Glutathione is usually maintained in a reduced state in the cytosol of cells

Margin Note 18.2
Biologically helpful reactive oxygen species

During the respiratory burst in phagocytic cells, free radicals such as superoxide are produced by the activation of a membrane-bound NADPH oxidase in an oxygen-dependent process. The superoxide radicals produced are released into a vacuole, helping to destroy its engulfed pathogens (*Chapter 4*).

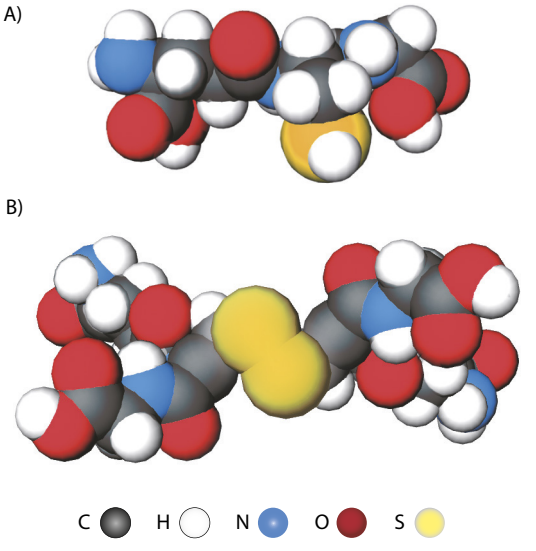


Figure 18.5 Computer drawn models of the (A) reduced and (B) oxidised forms of glutathione. See also *Chapter 13*.

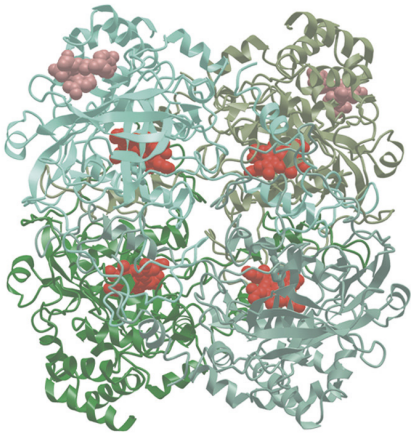


Figure 18.6 Molecular model of catalase. Its four haem groups and their associated Fe^{2+} are shown in red, with two nicotinamide phosphate coenzymes ($NADP^+$) indicated in light red. PDB file 1DGF.

and protects against free radical damage. The enzyme superoxide dismutase (see *Figure 10.26*) also removes superoxide radicals by converting them to hydrogen peroxide and dioxygen. Catalase (*Figure 18.6*) then catalyses the oxidation of the hydrogen peroxide to water and dioxygen:

Superoxide dismutase

$$2H^+ + 2O_2^{\bullet-} \rightarrow H_2O_2 + O_2$$

Catalase

$$2H_2O_2 \rightarrow 2H_2O + O_2$$

Glycated proteins

Both intra- and extracellular proteins are subject to posttranslational changes during ageing. For example, proteins exposed to reducing sugars, for example glucose or fructose, may undergo glycation. This occurs as a nonenzymatic reaction between free carbonyl groups of reducing sugars in their acyclic form and amino groups of the protein to form a Schiff base in a freely reversible reaction (*Figure 18.7*). The Schiff base is unstable and rearranges to a more stable Amadori product; this reaction is effectively irreversible. The extent of protein glycation *in vivo* depends on the degree and duration of blood hyperglycaemia. However, glycated proteins may undergo further reactions to form cross-linked fluorescent structures called advanced glycation end products (AGEs). Approximately 30 different AGEs have been identified. The most common AGE found *in vivo* is carboxymethyllysine; other examples are pyrraline and pentosidine (*Figure 18.8*).

Advanced glycation end products accumulate with age, particularly on structural proteins, such as collagen, which has a long biological half-life, and they can increase the cross-linking of individual proteins. Such changes have deleterious effects, since they decrease the elasticity and permeability of the extracellular matrix and impair the flow of nutrients into and waste products out of cells. Since a large proportion of the elderly population has diabetes or impaired glucose tolerance (*Chapter 7*), their hyperglycaemia means the harmful effects of glycation are more likely.

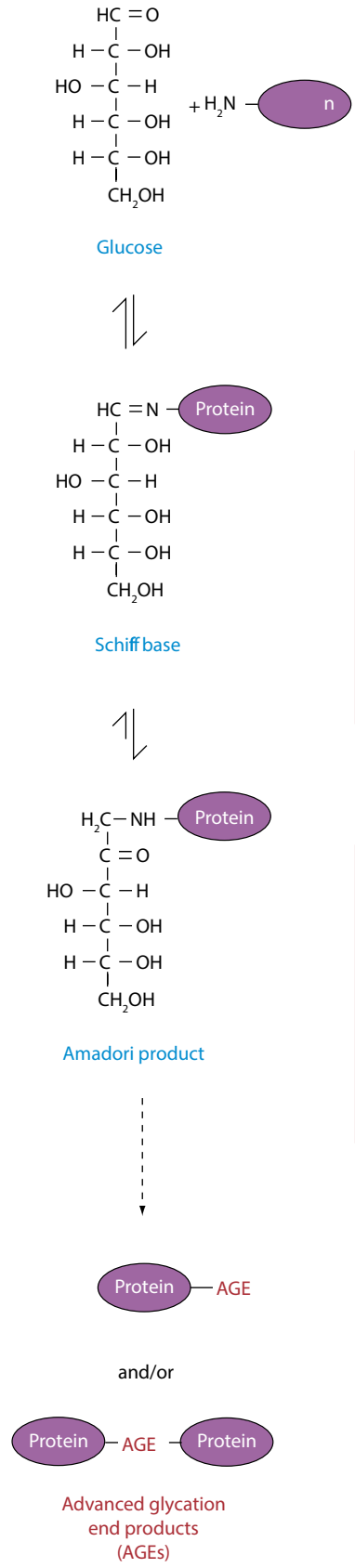


Figure 18.7 The reaction between glucose and a protein to form a glycated protein (Schiff base and Amadori product) and its subsequent conversion to an advanced glycation end product (AGE).

Margin Note 18.3 Glycation and glycosylation

Glycation should not be confused with glycosylation. While the former is a (relatively) straightforward chemical reaction, glycosylation involves the addition of sugar residues from nucleotide sugar donors to acceptor substrates in biochemical reactions catalysed by glycosyltransferases. See, for example, the formation of blood groups in *Figures 6.9–6.11*.

Margin Note 18.4 Glycated haemoglobin

The percentage of glycated haemoglobin in a diabetic patient should be determined regularly to assess therapy compliance over the previous six to eight weeks (*Figure 1.21* and *Chapter 7*). Diabetic patients with concentrations of glycated haemoglobin above a target value of 48 mmol/mol (6.5% of total haemoglobin) have poor blood glucose control, and their treatment or compliance requires reviewing.

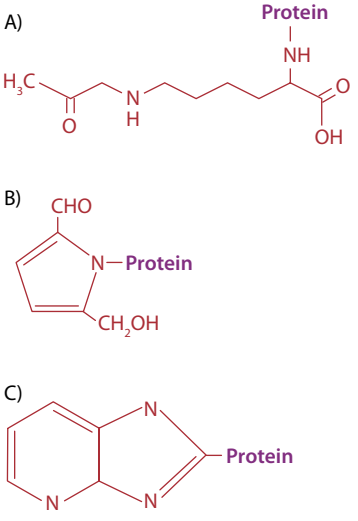


Figure 18.8 The advanced glycation end products (AGEs) (A) carboxymethyllysine, (B) pyrraline and (C) pentosidine.

Waste products

During ageing, increasing amounts of waste material accumulate in the cytoplasm of cells. Many of these are waste products of normal cellular metabolism. For example, lipofuscins are yellow-brown pigments (*Figure 18.9*) produced by degeneration of cell membranes and organelles, probably by the free radical peroxidation of membrane lipids. During ageing, lipofuscins accumulate in lysosome-like structures of many types of cells but particularly in nondividing cells, such as those of muscle and the central nervous system. Lipofuscin molecules are strongly cross-linked and chemically inert, and are not susceptible to digestion by lysosomal enzymes (*Chapter 16*). It has been suggested that a gradual accumulation of substances, like lipofuscins, in cells interferes with their normal function, although there is no conclusive evidence for this. Furthermore, there is no correlation between the amount of lipofuscin accumulated and the reduction in cell function and survival.

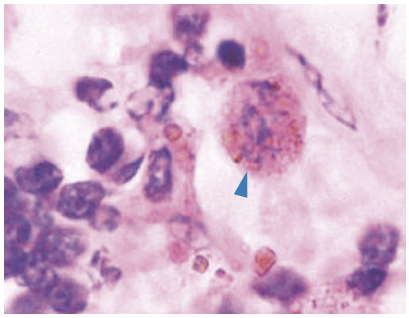


Figure 18.9 Light micrograph of a portion of a macrophage containing lipofuscin inclusions derived from lysosomes. A lipofuscin particle is indicated. Courtesy of Dr T. Caceci, Morphology Research Laboratory, Virginia-Maryland Regional College of Veterinary Medicine, USA.

The accumulation of abnormal proteins and the error-catastrophe theory

The error-catastrophe theory suggests that cellular dysfunction and, ultimately, cell death arises due to an accumulation of abnormal proteins. Protein biosynthesis involves the transcription of DNA to give mRNAs, which are transported to the cytoplasm where they are translated by ribosomes to form polypeptides (see *Figures 13.32* and *16.20*). Random mistakes in transcription and/or translation can lead to the formation of abnormal proteins, whose accumulation could impair cell functions. If the protein in question is an enzyme, such an error may lead to a malfunctioning enzyme and cellular dysfunction. Although enzyme activity is known to decline with ageing, it has not always been possible to demonstrate any changes in enzyme structure with age. It seems that proteins are biosynthesised appropriately in older cells and most subsequent changes to their structures occur posttranslationally. However, studies have indicated that some enzymes, for example liver aldolase, have altered conformations in older cells. This suggests that enzyme molecules retained inside the cell for long periods are slowly denatured and consequently lose their biological activity. In younger cells, the original conformations of proteins can be restored by cycles of denaturation followed by renaturation. Weak interactions that confer shape to the denatured form of protein molecules are broken, allowing them to fold back to their original shape. This process therefore corrects the defective shape of denatured enzymes and produces molecules that are as efficient as a newly synthesised enzyme. Unfortunately, such repair mechanisms lose their efficacy as cells age.

Genome-based theories

There is evidence to suggest that ageing is under genetic control. Several genetic-based theories to explain the effects of ageing have been suggested; two will be discussed here:

1. programmed ageing
2. gene mutations and ageing

Programmed ageing

The theory of programmed ageing suggests that each species has an in-built biological clock that genetically determines and restricts its maximum lifespan. In the 1960s, Leonard Hayflick (1928–2024) demonstrated that cells are restricted in the number of times they can enter the cell cycle (see *Figure 15.6*) by an inherent programme of senescence. He showed that cultured fibroblast cells derived from human embryos could undergo 50 cell divisions, whereas those from adults were limited to about 20. In culture, the number of divisions is constant for each type of cell, which is referred to as the **Hayflick limit**. Furthermore, the factors that control the number of divisions are intrinsic to the cell and are not influenced by their environment. For example, if the nucleus of an old cell is transplanted into a young one from which the original nucleus has been removed, the resulting cell has a lifespan that reflects that of the transplanted nucleus. Also, when cells grown in culture are frozen and then recultured, they appear to retain the memory of the number of times they have already divided in the original culture. Hence, they only complete the 'unused' number of cell divisions in accordance with their Hayflick limit. It therefore appears that there is a biological clock within all cells. This biological clock, at least in part, resides in the telomeres, which are extensions of DNA found at the ends of chromosomes (*Figure 18.10* and see *Box 17.1*). Telomeric DNA protects the ends of DNA molecules from damage. However, when DNA is duplicated prior to cell division, its telomeric DNA is not replicated. Thus, after each cell division the telomeres become shorter in length. Once the telomeres shorten to a particular length,

the cell can no longer divide and dies. Telomerase (*Figure 18.11*) activity can catalytically add telomeric repeats to the ends of chromosomal DNA (see *Figure 1.3*). Hence, it can prevent the shortening of telomeres and enable the cell to divide continuously. Most somatic cells contain an inactive form of telomerase, although some types, such as haemopoietic and cancer cells, have permanent telomerase activity and can divide indefinitely and are therefore potentially immortal (see *Box 4.5* and *Chapter 17*). Telomerase has both protein and RNA components. Mutations in *TERC*, a gene that encodes its RNA portion, decreases the activity of the enzyme and is associated with telomeres of reduced length and therefore with increased rates of ageing. Individuals with *TERC* mutations are believed to be biologically several years older than those of a similar calendar age but with wild-type forms of *TERC*.

The suggestion that ageing is genetically programmed has received some criticism. For example, the number of divisions occurring *in vitro* may be different from those that happen *in vivo*. Furthermore, some cells, such as cardiac muscle cells and neurons, do not divide after birth, and so programmed ageing may not apply to these cells, although many neurons do die during life resulting in the loss of brain tissue (*Section 18.2*).

Gene mutations and ageing

The gene mutation theory of ageing suggests that the number of mutations accumulated during life leads ultimately to cell, tissue and organ malfunctions and eventually death. Indeed, it is well established that mutated genes can alter the activities of cells (*Chapters 15* and *17*).

Genes are composed of DNA. The cell has several mechanisms to repair damaged, that is, mutated, DNA. Enzymes within the nucleus excise the damaged region of the gene and add back a new set of nucleotides using the undamaged DNA strand as a template. The gene mutation theory suggests that, with time, such DNA repair mechanisms become less efficient and some mutations are not repaired, leading to dysfunctional changes. In support of this theory, DNA obtained from the liver cells of older mice has been found to have a greater number of mutations than similar cells from younger ones. In addition, liver cells obtained from strains of mice with a short lifespan show greater numbers of mutations than similar cells from longer lived strains.

Radiation is known to cause mutations (and also stimulates the production of free radicals [see above]), leading to a shortening of the lifespan of cells. This has led to the suggestion that natural radiation may accelerate ageing processes.

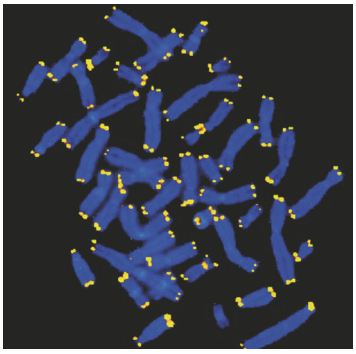


Figure 18.10 Telomeres of human chromosomes are stained to appear brighter than the rest of the chromosomes. Note these structural components of chromosomes are situated at the ends of the chromosomes. See also *Box 17.1*. Courtesy of Dr C. Counter, Department of Pharmacology and Cancer Biology, Duke University, USA.

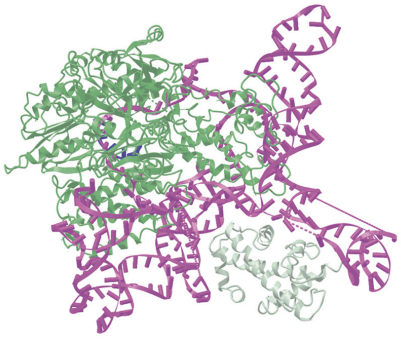


Figure 18.11 Molecular model of the catalytic core lobe of human telomerase (green) with a bound DNA substrate. Associated histones are depicted in the lighter green. PDB file 7BG9.

18.4 AGE-RELATED CLINICAL CONDITIONS

Several clinical conditions in particular show an increased incidence in elderly patients. These include:

- Alzheimer disease
- arthritis
- cancer
- cardiovascular disease
- cataracts
- Parkinson disease
- type 2 diabetes

Alzheimer disease

Alzheimer disease (AD) is a degenerative condition of the brain in which nerve cells, usually in the hippocampus, temporal and parietal regions, lose function and die. Late-onset AD causes approximately 50% of cases of dementia in the elderly, making it the most common cause of the condition in this group. In the UK, about 5–10% of the population over the age of 65 develops AD, and this increases to over 20% of those over the age of 80 years (*Figure 18.12*). Indeed, dementia and AD are the leading cause of death in the UK.

Alzheimer disease is characterised by the presence of extracellular proteinaceous plaques in the brain (*Figure 18.13*). The plaques are resistant to enzymatic and chemical digestion and remain in the brain tissue even after neuron death. They consist mainly of a core of β-amyloid peptides (Aβ) 40–42 amino acid residues long (*Figure 18.14*) entangled with tau proteins, and surround degenerating nerve terminals.

The Aβ peptides are formed by two specific hydrolyses of a β-amyloid precursor protein (APP) that are catalysed by β-secretase and γ-secretase (*Figure 18.16*). The APP is a cell surface receptor that is believed to have roles in synaptic transmission, memory and learning.

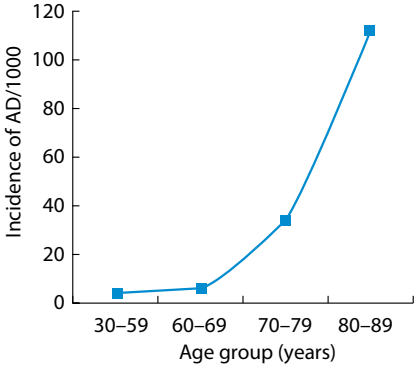


Figure 18.12 The increasing incidence of Alzheimer disease (AD) with age. Redrawn from The Association of the British Pharmaceutical Industry website.

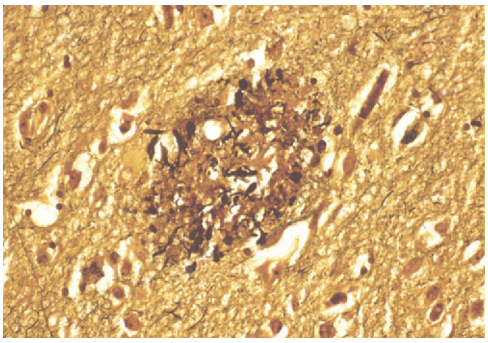


Figure 18.13 An amyloid plaque in the brain of a patient who died from Alzheimer disease. Courtesy of Alzheimer's Disease Education & Referral Center, National Institute on Ageing, USA.

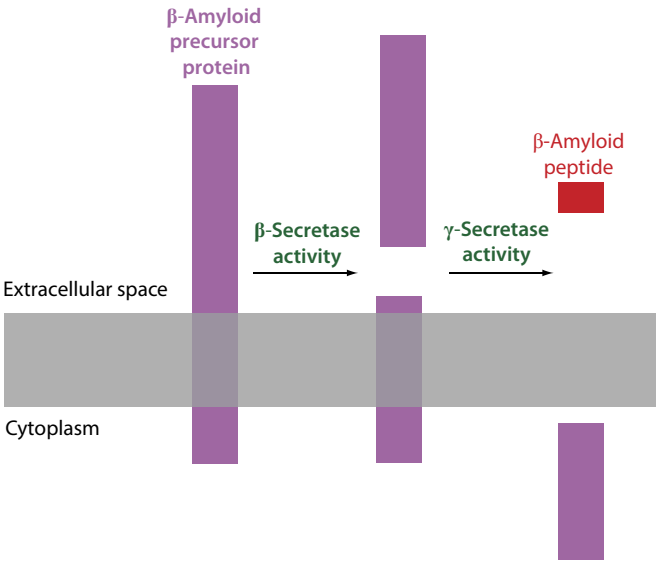


Figure 18.16 The release of β-amyloid peptide by the proteolytic activities of secretases. Redrawn from Haass C and Steiner H (2002) Alzheimer's disease – secretase: a complex story of GxGD-type presenilin proteases. *Trends in Cell Biol* **12**:556–562. With permission from Elsevier.

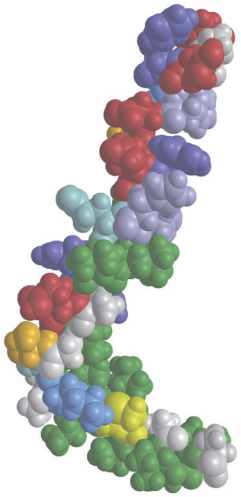


Figure 18.14 Molecular model of a β-amyloid peptide. PDB file 1IYT.

BOX 18.2 AMYLOID AND AMYLOIDOSES: A BIOCHEMICAL MISNOMER?

The names of amyloid plaques and the amyloid complexes that occur in a group of diseases called the amyloidoses (singular: amyloidosis) are, despite them being proteinaceous, named from their visual resemblance to starch (from the Greek *amylon* meaning starch and *eidos*, form).

Amyloidoses are rare diseases, which normally affect those in their late middle age or the elderly. They are characterised by the deposition of aberrant protein complexes in a variety of organs. Amyloid deposits can accumulate in multiple organs to the extent that the condition can be an immediately life-threatening disease and require emergency treatment. Over 30 different amyloidoses are known, each differing in the protein involved; hence, like prion diseases (*Chapters 2* and *15*), they are protein conformational diseases. Amyloidoses are named by stating the protein involved followed by the generalised term. The two most common are transthyretin (TTR) and amyloid light-chain (AL) amyloidoses. The latter was previously called primary amyloidosis. Transthyretin is a thyroxine and vitamin A transporter synthesised in the liver. In TTR amyloidosis, TTR molecules misfold and aggregate as amyloid complexes that disrupt the functions of, for example, the nervous system (NS), heart and gastrointestinal tract (GIT). Immunoglobulin AL amyloidosis is normally a plasma cell disorder involving the secretion of free unstable immunoglobulin light chains (*Chapter 4*). These can infiltrate and form amyloid deposits in the heart, kidneys, GIT, liver and NS. The major pathology is due to cardiac dysfunction.

Diagnoses of amyloidoses are made by taking biopsies of the suspected

Margin Note 18.5 Tau protein

Tau protein is expressed mainly in the brain, where it stabilises and orientates microtubules (MTs) necessary for the transport of materials along axons (*Figure 18.15*). The dephosphorylation of tau promotes the rapid and extensive polymerisation of MTs. In contrast, the phosphorylation of tau decreases its ability to promote MTs' assembly. In Alzheimer patient brains, tau is hyperphosphorylated, and its ability to stabilise MTs is impaired.

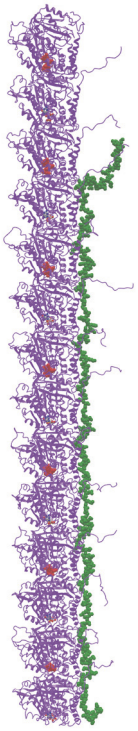


Figure 18.15 Molecular model of tau protein (green) bound to a portion of a microtubule. Guanosine di- and triphosphates and their associated Mg^{2+} are depicted in red. PDB file 7PQP.

organ(s). Treatment is aimed at alleviating any underlying chronic condition or disease. Immunoglobulin AL amyloidosis is treated with chemotherapy in a manner akin to that of cancer medication. Renal amyloidosis may require kidney dialysis and a kidney transplant. An autologous stem cell transplantation may be offered where suitable. Transthyretin amyloidosis is treated with TTR protein stabilisers, such as tafamidis, or by gene silencing therapies. Unlike AL amyloidosis, an organ transplantation (*Chapter 6*) is less of an option, since it would not address the other organs damaged by the condition. Advances in methods of diagnosis and therapies have, however, greatly changed the prognosis for patients suffering amyloidoses.

Approximately 5–10% of AD cases are familial, that is, are inherited forms of the condition, with most cases being sporadic or senile AD, the risk of which increases with age. Familial or early-onset AD is associated with mutations in two genes: *PSEN1*, located at 14q24.2, which encodes presenilin-1, the catalytic component of γ-secretase; and *PSEN2*, located at 1q42.13, which is a homologue of *PSEN1*. Mutations in *PSEN1* are more common and appear to cause more aggressive forms of AD with, in some cases, onset occurring before the age of 30, although 45–60 years is more likely for early-onset AD. Sporadic cases of AD are more likely to occur in people with the gene for the apoEε4 variant of apolipoprotein E (*Chapter 14*), especially if they are homozygous. Environmental risk factors, including exposure to aluminium, head injuries and some viral infections, are also associated with AD. Alzheimer disease is also associated with a decline in choline acetyltransferase activity (*Figure 18.17*), which is required for the biosynthesis of acetylcholine from choline and acetyl CoA:

$$(CH_3)_3{-}N^+{-}(CH_2)_2{-}OH + CH_3CO{-}SCoA \xrightarrow{\text{Choline acetyltransferase}} (CH_3)_3{-}N^+{-}(CH_2)_2{-}O{-}CO{-}CH_3 + CoASH$$

choline — acetyl CoA — acetylcholine — coenzyme A

Indeed, there is a correlation between reduced enzyme activity, on the one hand, and the number of plaques and severity of the dementia, on the other.

The onset of AD is insidious, starting with periods of forgetfulness, leading to a confused state and eventually frank dementia. It can be divided into three stages. The first stage may last two to four years and is associated with memory loss, personality changes and disorientation concerning times and dates. The memory loss in this first stage is difficult to distinguish from the normal forgetfulness that occurs in the elderly. The second stage may last for several years and includes confusion, depression, inappropriate social behaviour, agitation and an inability to carry out the activities associated with daily living. The memory lapses are more frequent, and patients often forget what they were doing only a few minutes previously. During this stage, personal hygiene is often neglected and vocal communication becomes impaired, as the patient has difficulty in remembering words. The final stage lasts for one to two years, although it can extend for as long as 10 years. During this stage, affected individuals fail to recognise their family, suffer from urinary and faecal incontinences, and cannot communicate. The affected individuals are usually institutionalised at this stage. Once severe dementia develops, life expectancy is about two or three years. The general clinical course of AD is about eight years, with patients often dying from infections such as pneumonia, accidents and occasionally respiratory arrest.

There is no simple and completely accurate test to help diagnose AD. Diagnosis is based on clinical tests supported by imaging techniques. A medical history of the suspected AD patient is taken followed by cognitive tests to evaluate memory recall, the ability to concentrate and communicate, and general awareness. Blood tests need to be conducted to assess liver, thyroid and kidney functions, HbA_{1c} to discount diabetes, and vitamin B_{12} and folate status to exclude other causes of symptoms associated with dementia. A lumbar puncture may be necessary to obtain cerebrospinal fluid (CSF) in which β-amyloid and tau concentrations are determined, given that both accumulate in the brains of patients with AD. However, a firm diagnosis of AD can only be given by a histological examination of brain tissue after death. However, brain imaging techniques, particularly magnetic resonance imaging (MRI) and positron emission tomography (PET), have proved increasingly useful in diagnosis. Magnetic resonance imaging allows the structure of the brain to be studied in a noninvasive manner. The technique is based on the principle of nuclear magnetic resonance and uses powerful magnetic

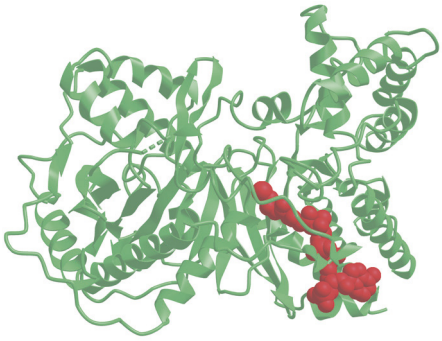


Figure 18.17 Molecular model of choline acetyltransferase. A bound coenzyme A is depicted in red. PDB file 2FY4. ⏎

fields to obtain chemical and physical information about the molecules within the brain. A computer then uses this information to generate an image of the internal structure of the brain. Physical lesions to the brains of AD patients are clearly visible using MRI (*Figure 18.18*).

Positron emission tomography allows the activities in different parts of the brain to be estimated. This is achieved by adding labelled glucose or water to the blood and then monitoring the flow of blood through the brain or the rate of glucose metabolism in its different parts. Again, a computer can analyse this information to produce digital images that highlight differences in the activities of the brains of normal and AD patients (*Figure 18.19*).

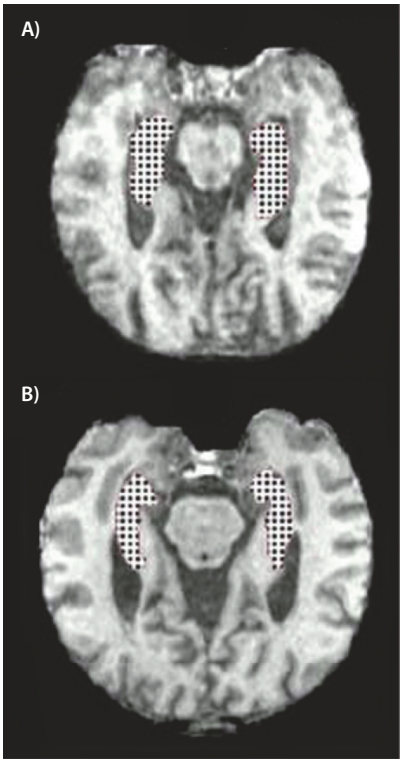


Figure 18.18 Magnetic resonance imaging (MRI) scans of (A) a normal brain and (B) a brain showing atrophy in the hippocampal area from an Alzheimer disease patient. Courtesy of Dr M. de Leon, New York University School of Medicine, USA. ⏎

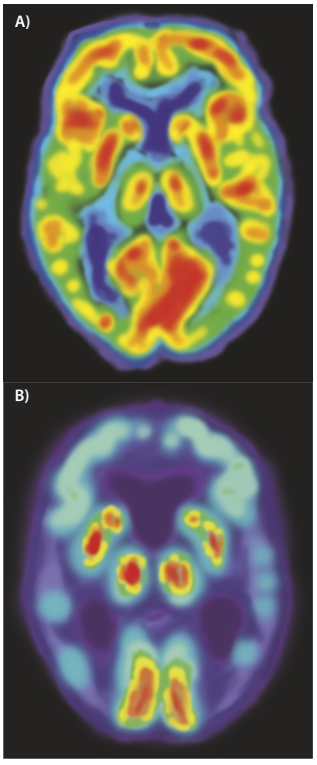


Figure 18.19 Positron emission tomography (PET) images showing activity of (A) a normal brain and (B) the brain of an Alzheimer disease patient. Courtesy of Alzheimer's Disease Education & Referral Center, National Institute on Ageing, USA. ⏎

Given the difficulties in diagnosis, cases of AD are under-reported. Early diagnosis is of considerable benefit, since it would allow all concerned to make informed early medical, social and legal decisions about treatment and care for a patient. An early diagnosis would allow drug treatment and care that could delay institutionalisation and substantially reduce costs. Conversely, an early test that indicated an absence of AD in suspected cases would alleviate the uncertainty and anxiety faced by the patients and their families. It has been suggested that a plasma variant of tau that is phosphorylated at threonine217 (p-tau217), which is associated with AD, could be the basis of a blood-based biomarker test that is as accurate as lumbar punctures or PET scans in detecting and monitoring the progression of AD. The test involves assessing and demonstrating increased concentrations of p-tau217 in the blood. It is believed that this test will allow an early diagnosis of AD, allowing earlier treatment than at present, benefiting the patient by slowing down the progression of AD. However, even if a quick diagnosis were possible, there is no effective treatment for AD. Sufferers may be placed on medication to alleviate symptoms such as depression and anxiety. Drugs that inhibit the degradation of acetylcholine within synapses, such as acetylcholinesterase inhibitors, are used in treatment. These drugs can delay the impairment of cognition, behaviour and functional abilities. Vitamin E (*Chapter 10*) treatment has also been used, although some studies have suggested it is of little benefit. In 2023, the United States Food and Drug Administration (FDA) approved

the clinical use of the drug lecanemab to treat AD. Lecanemab is a human antimonoclonal antibody (see *Box 4.5*) that can reverse amyloid plaque pathology and thereby slow down the progression of AD. The drug is administered as an intravenous infusion every two weeks, and has been shown to reduce cognitive decline in patients in the early stages of AD.

Arthritis

Arthritis is an inflammation of the joints. Osteoarthritis affects the joint cartilage and underlying bone, producing swelling, pain and restricted movement. It is particularly associated with increasing age, although it can occur in younger individuals who excessively use their joints in work or athletic activities. Osteoarthritis attacks fingers, hip joints and knees (*Figure 18.20*) but, unlike rheumatoid arthritis (*Chapter 5*), affects each patient differently and does not always cause pain and inflammation. X-raying of joints usually shows some degree of osteoarthritis in nearly all elderly people, although few present with any symptoms. In severe cases, the joints of the fingers often show overgrowth referred to as Heberden nodes (*Figure 18.20*), although these tend not to be painful. Osteoarthritis cannot be cured, although mild exercise can improve joint mobility.

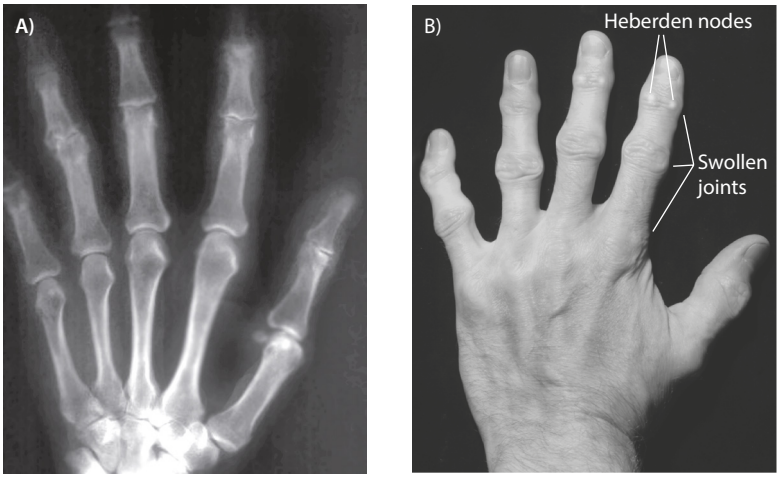


Figure 18.20 (A) X-ray image of a middle-aged female patient with osteoarthritis of the left hand who presented with pain and swelling in the finger joints. (B) The left hand of a sufferer of osteoarthritis. Courtesy of Dr P. Young, Department of Radiology, University Hospitals of Cleveland, USA.

Rheumatoid arthritis is characterised by a chronic inflammation of the joints that usually arises from an autoimmune reaction (*Chapter 5*). It is also more common in the elderly, although its onset can occur in any age group. The result is severe pain and disability. What initiates rheumatoid arthritis is unclear, although a variety of pathogenic bacteria, especially mycobacteria (*Chapter 2*), have been implicated. The treatment of rheumatoid arthritis involves using nonsteroidal anti-inflammatory drugs. The surgical replacement of hip or knee joints may also be required in patients who become severely disabled.

Cancer

In general, the incidence of most cancers increases with age, mainly due to decreases in the body's abilities to repair mutated genes and to a weakening of the immune system with a corresponding reduction in its ability to kill cancerous cells. The incidences of cancers increase up to 90 years of age and are one of the major causes of death (*Figure 18.21*). The number of new cases of cancer is projected to exceed 35 million by 2050.

Two main hypotheses have been proposed to explain the link between cancer and ageing. First, an age-related accumulation of carcinogenic substances may increase the incidence of cancers in the elderly. This process is independent of the senescence changes that occur in the ageing body and described in *Section 18.2*. The second hypothesis is that age-related changes in cells make them more vulnerable to becoming cancerous. Changes in immune, nutritional, metabolic and endocrine status occur with

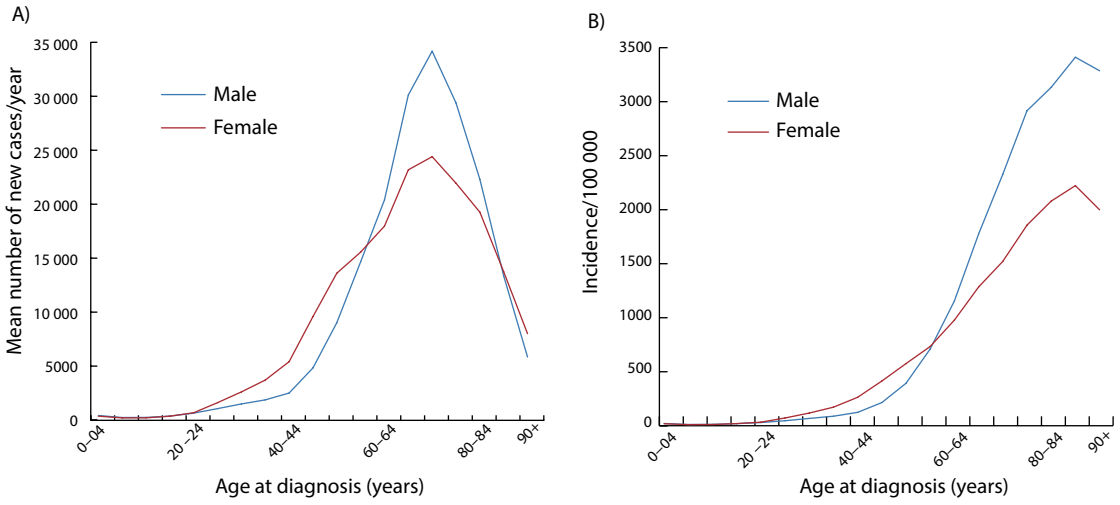


Figure 18.21 (A) The number of new cases of all cancers other than nonmelanoma skin type and (B) the increasing incidence of all cancers, again other than nonmelanoma skin type, with age. Data courtesy of Cancer Research UK. ⏎

age and may create a more favourable environment for the induction of cancer. Such physiological changes may decrease the effectiveness of protective cell processes, such as the detoxification of mutagenic agents (*Chapter 12*) or the repair of damaged DNA (*Chapter 17*).

Cardiovascular disease

Many of the changes in the cardiovascular system may be caused by disease rather than old age per se (*Chapter 14*). The concentration of cholesterol in the plasma increases with age, and over the years is thought to contribute to the high incidence of mortality from coronary heart disease. This is especially so if other risk factors known to be associated with atherosclerosis, for example eating an inappropriate diet, tobacco smoking and a lack of exercise, are present. However, the risk of coronary heart disease can be decreased by appropriate reversal of these inappropriate lifestyle choices.

Cataracts

A cataract is a partial or complete opacification of the lens of the eye that causes blurred vision. This affects the passage of light through the lens, causing blindness. There are many different types of cataracts but one of the most common is senile cataract, where ageing leads to denaturation of lens crystallins (see *Figure 18.2*). Diabetes is also a noted risk factor for cataracts (*Chapter 7*). Cataracts are treated by the removal of the opaque lens proteins and the insertion of a replacement intraocular plastic lens into the lens capsule (*Figure 18.22*).

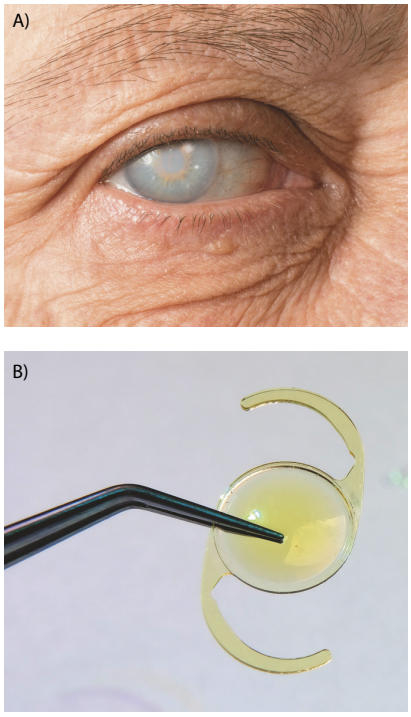


Figure 18.22 (A) An eye with a prominent cataract. (B) An elastic intraocular lens used in cataract replacement surgery. (A), courtesy Shuutterstock.com/Sruilk; (B), courtesy of Shutterstock.com/Mohammed_Al_Ali. ⏎

Parkinson disease

The major defect in Parkinson disease (PD) is degeneration of dopamine-secreting nerve cells (see *Chapter 7*), although other neurons and neurotransmitters may also be affected. Parkinson disease is distinct from AD (see above) in that different nerve cells are affected and there is a loss of motor function, which is usually unaffected in AD. A further feature of PD is the presence of cytoplasmic proteinaceous inclusions called Lewy bodies (*Margin Note 18.6*) in some of the surviving neurons. Some researchers believe that an excess of free radicals causes the neuron degeneration associated with PD. The condition affects 1–2% of people over the age of 70 years. Patients have severe attacks of tremors that affect one hand and then spread to the leg on the same side and then to other limbs. The average survival time for a Parkinson patient is 8–10 years following diagnosis.

Type 2 diabetes mellitus

Many older people have some degree of impaired glucose tolerance that can be severe enough to be classified as type 2 diabetes mellitus, as described in *Chapter 7*. The main reason for high concentrations of blood glucose in the elderly is an increased resistance

to the stimulatory effects of insulin in peripheral tissues, which is associated with increased concentrations of insulin after meals. Diabetes in older people is strongly influenced by diet and exercise; improvements in both can provide better glycaemic control and extend life expectancy.

18.5 ENERGY RESTRICTION AND AGEING

A reduced energy intake ('calorie restriction') is known to slow the rate of ageing and the onset of age-related disorders, such as cancer (breast, lymphomas and prostate), nephropathy, cataracts, diabetes, hypertension, hyperlipidaemia and autoimmune diseases. This has been demonstrated in a variety of species, including chickens and rodents, and is also believed to be true for humans. The effects of calorie restriction were demonstrated in the 1930s using laboratory rats. One group of rats was allowed to feed freely while another was fed a diet containing only 30% of the calories of the first group, although they were provided with sufficient protein, fats, vitamins and minerals to maintain normal health. The calorie-restricted rats lived for four years compared with only three years for those allowed to feed freely. In addition, the energy-restricted rats developed fewer age-related diseases.

Studies on energy restriction have been performed in primates with encouraging results. Long-term studies on Rhesus monkeys showed that calorie restriction reduced the incidence of heart disease, diabetes and hypertension, and was associated with a decreased concentration of blood cholesterol. Energy restriction is, however, difficult to apply to humans because many people are unable to reduce their calorie intakes by appreciable amounts for the extended periods required. However, it may be possible to motivate people to do this, especially those with family histories of age-related diseases such as cancer and neurodegenerative disorders.

The mechanism by which energy restriction increases the lifespan is unclear but studies have shown that it is associated with a reduction in age-associated mutations when compared with normal diets. This was demonstrated by examining mutations in mice lymphocytes at four weeks, six months and one year of age.

A high-energy diet may increase free radical-mediated damage, as the increased availability of nutrients to mitochondria increases the production of superoxide radicals. Thus, a calorie-restricted diet appears to reduce free radical damage to lipids, proteins and DNA and improves the general antioxidant status. Calorie restriction in animals has also been shown to reduce levels of tissue AGEs. The benefits of energy restriction, however, depend on preventing malnutrition but reducing the overall calorie intake rather than that of any particular nutrient.

BOX 18.3 THE YOUNG AND OLD OF AGEING: PROGERIA OR HUTCHINSON GILFORD SYNDROME

Progeria (from the Latin and Greek words *pro* and *geraios* meaning early and old age, respectively) or Hutchinson–Gilford syndrome is a disorder that causes premature ageing. The syndrome was first described by two British surgeons, Jonathan Hutchinson (1828–1913) in 1886 and Hastings Gilford (1861–1941) in 1897. Progeria is a rare condition affecting one in four million people worldwide. About 400 new cases arose in 2023. The disease is not restricted to any ethnic group or geographical area but males are affected one and a half times more frequently than females.

Children with progeria age about 10 times faster than normal; thus, a child of 8–10 years old looks like an 80 year old (*Figure 18.24*).

Progeria is caused by an autosomal heterozygous mutation (*Chapter 15*). A single base change in the lamin A gene, *LMNA*, found at 1q22, leads to aberrant RNA splicing and the deletion of 50 amino acid residues (Δ50) producing a truncated and toxic nuclear lamin A precursor, sometimes called progerin (see *Figure 13.32* for an outline of splicing). In most cases, the mutation is undoubtedly

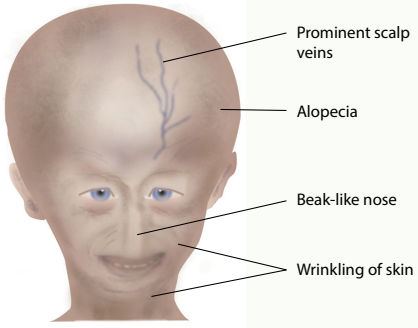


Figure 18.24 Schematic of the head of a progerian child.

Margin Note 18.6
Lewy bodies

Lewy bodies (*Figure 18.23*) are named after the German neurologist Friedrich Lewy (1885–1950). They consist of concentric spherical aggregates of α-synuclein, a presynaptic protein whose function(s) is poorly understood, and ubiquitin. Lewy bodies are found in vacuoles within neurons of the midbrain and brainstem in patients with Parkinson and Alzheimer diseases (PD and AD) and Lewy body dementia (LBD). The latter (LBD) covers two clinical entities, that is, dementia with Lewy bodies and PD dementia. It is thought to be underdiagnosed and, although poorly understood clinically, its features overlap with those of PD and AD, although the deficiency in acetylcholine is more severe.

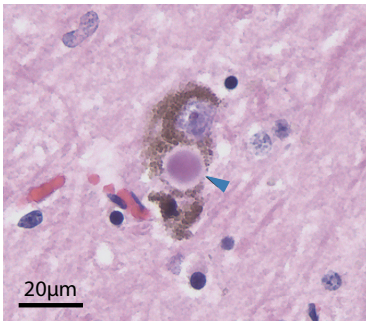


Figure 18.23 Photomicrograph of a stained Lewy body (indicated with an arrowhead) in a neuron of the substantia nigra of the brain, obtained from a patient who sadly died from Parkinson disease (PD). The brown/black material is neuromelanin that is ordinarily present in neurons of this region of the brain. Courtesy of Wikimedia/Tulemo.

de novo and is not found in either parent. Given that lamin A is an essential protein of the nuclear skeleton, the result is nuclear deformation and dysfunction, and development of the syndrome. Interestingly, other mutations in *LMNA* and related genes produce a variety of somewhat less severe clinical conditions but with features that overlap those of progeria. These include mandibuloacral dysplasia, dilated cardiomyopathy with hypergonadotropic hypogonadism and Werner syndrome.

The development and appearance of children with progeria are seemingly normal in the first two years of life, after which the characteristic ageing changes take place with a rapidity that can be shocking. Their features are remarkably similar and include thinning and wrinkling of skin, prominent scalp veins, loss of subcutaneous fat, alopecia (loss of hair), a beak-like nose, short stature, thin limbs with stiff swollen joints, severe arthritis, osteoporosis and high-pitched (squeaky) voice, but they have normal or high intelligence. Some features of ageing are, however, absent. For example, progerian children often present with delayed development of teeth and sexual maturity, but show no increases in the incidences of cancers, diabetes or cataracts. Fibroblasts isolated from patients with progeria and grown in tissue culture have a shorter lifespan than those from normal individuals of a similar age. This is due to the chromosomes of progeria sufferers having short telomeres (see *Figure 18.10*), which results in cells having a lower Hayflick limit. Furthermore, some studies have shown a decrease in the ability of cells from patients with progeria to repair damaged DNA, although this finding has not been supported by other studies. The life expectancy of progerian patients is 14–15 years (range of 6–27). Death usually occurs from a heart attack or cerebrovascular disease (*Chapter 14*).

A diagnosis of progeria is made on clinical grounds but can be difficult due to the rarity of the condition and its insidious onset in the early stages. However, the identification of the Δ50 mutation has enabled a diagnostic genetic test to be developed that allows an earlier identification of progeria in symptomatic children. Given its rarity, children with progeria often think they are the only ones with this disorder. At school they perform well and usually have a cheerful and open nature. However, because of their appearance, they are usually stared at by strangers and must learn to cope with such social problems from an early age.

There is no cure for progeria, but the drug lonafarnib (*Figure 18.25*) received FDA approval in 2020. It is a farnesyltransferase inhibitor that blocks the farnesylation of progerin and thus inhibits its build-up. Lonafarnib has been shown to increase the survival rate for progeria children by two and a half years. Children treated with lonafarnib show an increased flexibility in their blood vessels, improved hearing and bone structure, and gain weight. Patients may also be placed on a statin and low-dose aspirin therapy to delay symptoms of atherosclerosis (*Chapter 13*).

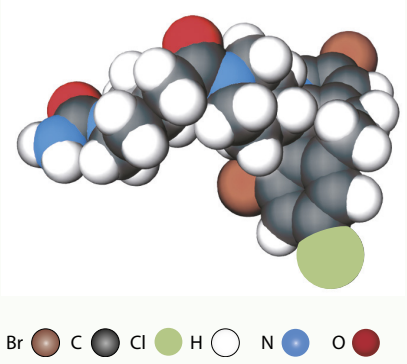


Figure 18.25 Computer drawn model of lonafarnib.

18.6 INVESTIGATING AND MANAGING CLINICAL CONDITIONS IN THE ELDERLY

Investigating and managing clinical conditions in the elderly involves a number of problems. Many conditions are more common in the elderly, and often the presentation of some of these diseases may differ from that in younger people. For example, diabetes mellitus in the elderly often presents as a complication of, for example, renal failure or impaired healing of wounds, instead of the classical signs of polydipsia or polyuria (*Chapters 7* and *8*) first seen in younger patients. Elderly people, particularly those with poor mobility, may also suffer from poor nutrition. Furthermore, they are often on multiple medications that may affect test results. The high incidence of many diseases in the elderly population justifies screening programmes to increase the chances of detecting such conditions at a more treatable stage. Clinicians, hospitals and geriatric clinics can all carry out screening. Some of the common investigations are listed in *Table 18.1*.

Investigation	Abnormality detected
Plasma [creatinine]	Renal impairment
Plasma [calcium]	Hyperparathyroidism/osteomalacia
Plasma [glucose]	Diabetes mellitus
Thyroid hormones	Hypo- and hyperthyroidism
Faecal occult blood	Large bowel carcinoma
Blood pressure	Hypertension
Cholesterol	Coronary heart disease
Mammography	Breast cancer

Table 18.1 Common clinical tests used in elderly patients.

18.7 CASE STUDIES

CASE STUDY 18.1

Joyce, a 70-year-old woman, presented with an ulcer on the sole of her right foot. This was causing her considerable pain and on examination her foot was cold and appeared ischaemic. The doctor suspected Joyce might be diabetic but she did not complain of thirst or polyuria. Random blood and urine specimens were taken and tested for glucose. The following results were obtained (reference range given in parentheses):

Urine [glucose]	Positive	(Negative)
Plasma [glucose]	16	(3.0–5.5 mmol/L)

Questions

(a) Does Joyce suffer from diabetes? (Consulting *Chapter 7* may be helpful.)
(b) If so, what type is it likely to be?

CASE STUDY 18.2

Harry, a 64-year-old priest, was noted for making mistakes while delivering his sermons. His wife noticed that these were becoming more common and that he was suffering from increasing lapses of memory. Harry was persuaded to visit his doctor and during his examination he appeared to be fully alert and held a normal conversation. However, in a word memory test, after five minutes he was able to recall only one out of every three words. He was referred for a more detailed neurological examination that was largely normal, although an MRI scan of his brain showed some possible hippocampal atrophy. No other significant changes were noted.

Questions

(a) What is a possible diagnosis?
(b) What advice should Harry and his family be given?
(c) How should Harry be treated?

CASE STUDY 18.3

Carla is a 12-year-old girl who showed normal development until about 18 months of age, when her parents started to notice some abnormal changes occurring, which unfortunately have progressed. She is of normal intelligence and her parents were not aware of any other family members showing similar features. Carla presents the following clinical features: progressive coarsening and wrinkling of skin, alopecia and a short stature. She also appears malnourished and therefore shows a general failure to thrive. Her biochemical profile is unremarkable except that the concentration of serum cholesterol is above reference values. Her urine was examined for hyaluronic acid, which was also found to be present in abnormally high concentrations. Based on clinical presentation and the presence of hyaluronic acid in her urine, Carla was provisionally diagnosed with progeria.

Questions

(a) How might the initial diagnosis of progeria be confirmed?
(b) If the diagnosis of progeria is confirmed, how should Carla be treated?
(c) Given a positive confirmation, suggest a likely prognosis for Carla.

18.8 SUMMARY

The deleterious effects of ageing processes are numerous and diverse. They affect cells, tissues, organs and systems. Ageing may be caused by a combination of 'wear and tear' with the accumulation of harmful metabolic products, damaged proteins, genetic mutations and the inherent ageing process that can be demonstrated in cultured cells. A number of diseases are associated with increasing age, including cardiovascular disease, type 2 diabetes and Alzheimer and Parkinson diseases.

There is some evidence that antioxidants in the diet may delay the onset of ageing processes. Energy-restricted diets have been shown to increase the lifespans of animals. Ideally, elderly patients should be routinely screened for diseases associated with old age.

QUESTIONS

1. Which of the following enzymes is an established intracellular antioxidant?
(a) Lactate dehydrogenase.
(b) Phenylalanine hydroxylase.
(c) Superoxide dismutase.
(d) Sucrase.
(e) γ-Secretase.

2. Which of the following statements is the correct definition of the term 'lifespan'?
(a) The mean age of individuals in a population.
(b) The minimum age of individuals in a population.
(c) The age when individuals in a population reach puberty.
(d) The maximum age that can be attained by members of a population.
(e) The maximum age that can be attained by most members of a population.

3. Which ONE of the following terms best describes a free radical?
(a) Molecules with a positive charge.
(b) Enzymes.
(c) Cross-linked proteins.
(d) Atoms or molecules with an unpaired electron.
(e) Ions with a negative charge.

4. In less than 100 words, outline the role of protein glycation in ageing.

5. List FOUR metabolic or molecular changes you might reasonably expect to see in ageing cells.

6. List FIVE clinical disorders that show an increased incidence with age.

7. Which of the following is NOT a characteristic feature of progeria?
(a) Accelerated ageing.
(b) Increased incidence of cancers.
(c) Osteoporosis.
(d) Alopecia.
(e) Short stature.

8. Complete *Table 18.2*.

Feature	Familial AD	Senile AD
Age of onset (years)		
Percentage of cases		
Associated gene or risk factor		

Table 18.2 For completion. ⏎

9. Which of the following statements are true?
(a) Many major diseases of the developed world, such as coronary heart disease, cancer and diabetes type 2, are age-related.
(b) Lifespan is the mean length of life of individuals in a population.
(c) Age-related changes to body organs include a 20–30% reduction in the weight and volume of kidneys as scar tissue replaces glomeruli.

(d) Glutathione is a tetrapeptide found in most cells, which can occur in reduced and oxidised forms because it contains a thiol (-SH) group.
(e) Only reducing sugars, for example glucose or fructose, that contain a Schiff base can glycate proteins.
(f) Cells grown in culture can only enter a restricted number of cell cycles, referred to as the Hayflick limit, which is genetically and environmentally controlled.
(g) Alzheimer disease is characterised by the presence of extracellular proteinaceous plaques containing mainly β-amyloid peptides (Aβ), about 60–62 amino acid residues long, and tau proteins.
(h) Osteoarthritis affects mainly the fingers, hips and knees joints, but is associated with less inflammation and pain than rheumatoid arthritis.
(i) Parkinson disease is associated with degeneration of neurons, which leads to severe tremors affecting the hands and legs.
(j) Progeria results from a mutation in *LMNA* at 14q24.2 that leads to the production of a shortened nuclear lamin A precursor called progerin.

FURTHER READING

Abramoff B and Caldera FE (2020) Osteoarthritis: pathology, diagnosis, and treatment options. *Med Clin North Am* **104**:293–311.

Abud T *et al*. (2022) Determinants of healthy ageing: a systematic review of contemporary literature. *Aging Clin Exp Res* **34**:1215–1223.

Bloem BR *et al*. (2021) Parkinson's disease. *Lancet* **397**:2284–2303.

Breijyeh Z and Karaman R (2020) Comprehensive review on Alzheimer's disease: causes and treatment. *Molecules* **25**:5789. doi: 10.3390/molecules25245789.

Calabrò M *et al*. (2020) The biological pathways of Alzheimer disease: a review. *AIMS Neurosci* **8**:86–132.

Cicinelli MV *et al*. (2023) Cataracts. *Lancet* **401**:377–389.

Cisneros B *et al*. (2023) Hutchinson-Gilford progeria syndrome: cellular mechanisms and therapeutic perspectives. *Arch Med Res* **54**:102837. doi: 10.1016/j.arcmed.2023.06.002.

Cummings JL (2004) Alzheimer's disease. *N Engl J Med* **351**:56–67.

da Costa JP *et al*. (2016) A synopsis on aging-theories, mechanisms and future prospects. *Ageing Res Rev* **29**:90–112.

Knopman DS *et al* (2021) Alzheimer disease. *Nat Rev Dis Primers* **7**:33. doi: 10.1038/s41572-021-00269-y.

Macicior J *et al*. (2021) Small-molecule therapeutic perspectives for the treatment of progeria. *Int J Mol Sci* **22**:7190. doi: 10.3390/ijms22137190.

Maruszak A *et al*. (2023) Predicting progression to Alzheimer's disease with human hippocampal progenitors exposed to serum. *Brain* **146**:2045–2058.

Park DC and Festini SB (2017) Theories of memory and aging: a look at the past and a glimpse of the future. *J Gerontol B Psychol Sci Soc Sci* **72**:82–90.

Taylor J-P *et al*. (2020) New evidence on the management of Lewy body dementia. *Lancet Neurol* **19**:157–169.

Taylor MS *et al*. (2022) Current approaches to the diagnosis and management of amyloidosis. *Intern Med J* **52**:2046–2067.

Taylor & Francis
Taylor & Francis Group
http://taylorandfrancis.com

Index

Note: Entries which are simply page numbers refer to the main text. Other entries have the following abbreviations immediately after the page number: B, Box; C, Case Study; F, Figure; M, Margin Note; T, Table.

A

AABB *see* Association for the Advancement of Blood and Biotherapies
A2 chromosomes, 465F, 466T
AAAs *see* abdominal aortic aneurysms
AAV *see* adeno-associated virus
ABC *see* ATP-binding cassette; ATP-binding cassette proteins
ABCD1 gene mutations, 499–500
abdominal aortic aneurysms (AAAs), 445M
ABL gene mutations, 522, 562
abnormal heart rhythms, 430–432
  see also arrhythmias
abnormal proteins, 574
ABO blood group, 152, 153F, 162–165, 186–187, 188C
  allommunisation, 172
  antibodies, 164–165
  antigens, 162–163
  discovery, 162M
  and ethnicity, 163T
  genotypes/phenotypes, 164T
  HDN, 170M
  HTR, 171
  incompatibility, 165
  laboratory tests, 168
abscesses (boils), 54–55
absorption, nutrients, 328–335
  carbohydrates, 328
  lipids, 330
  minerals and trace elements, 332–333
  nucleosides and nucleotides, 329, 330F
  peptides and amino acids, 328, 329F
  vitamins, 331–332
  water, 333–335
AC *see* adenylate cyclase
accuracy, 15
ACD *see* anaemia of chronic disease
ACE *see* angiotensin-converting enzyme
acetaminophen *see* paracetamol
acetoacetate, 210, 265F
acetone, 210, 235C, 265F
acetyl CoA, 370
N-acetyl-*p*-benzoquinoneimine (NAPQI), 364–365, 364F, 365F
N-acetylcysteine, 367F
acetyldehyde *see* ethanal
N-acetylglucosamine, 125
N-acetylglucosaminetransferase, 163–164
acetylsalicylate *see* aspirin
achlorhydria, 399
Achromobacter xylosoxidans, 85B
acid–base balance disorders, 263, 268–273
  definitions, 268
  diagnosis and treatment, 274
  metabolic acid–base disorders, 269–270
  mixed acid–base disorders, 272–273
  respiratory acid–base disorders, 270–272
  types and compensatory mechanisms, 268, 269T
acidosis, 253F, 263, 268, 269T
  see also metabolic acidosis
acids, digestion, 322, 322B, 326–327B
ACKR1 *see* atypical chemokine receptor 1
acne, 54, 55F, 235C
acquired immunity, 110
acquired immunodeficiency syndrome (AIDS), 2–3, 45, 58, 62–64B, 129, 130T, 172, 395, 408, 445, 529–530
  cryptococcosis, 46
  see also human immunodeficiency virus
acquired mitochondrial DNA defects, 498B
acridine orange, 179
acrocentric chromosomes, 465F, 472F
acromegaly, 205, 212, 436–437
ACTH *see* adrenocorticotrophic hormone
actin monomers, 443, 444F
actin-myosin contractile fibers, 503F
  see also muscle contraction
Actinomyces spp., 58–59
actinomycosis, 59
activated sludge (sewage) treatment plants, 23, 24F
activation coefficients, 304–305, 305F
active insulin, 211F
acute chest syndrome, 404
acute conditions, 3
acute gastritis, 339
acute graft rejection, 175, 176–177
acute heart failure, 434
acute hepatic necrosis, 365F
acute hepatitis, 350–351
acute inflammation, 107B
acute intravascular haemolysis, 165, 171
acute kidney injury (AKI), 243–244, 255–256
acute lymphoblastic leukaemia (ALLs), 5, 522, 559–560
acute myeloid (myeloblastic) leukaemia (AML), 188C, 495, 520B, 539T, 559–561
acute necrotizing ulcerative gingivitis (ANUG), 337
acute pancreatitis, 348
acute phase proteins, 107B

acute phase response, 99, 107B
acute promyelocytic leukaemia (APL), 471–472B, 561M
acute pulmonary oedema, 434
acute respiratory distress syndrome (ARDS), 32
acyclovir, 54, 54B, 54F, 75T, 100M, 337
ADA *see* adenosine deaminase
adalimumab, 121B
ADCs *see* antibody–drug conjugates
addiction, 373
  see also alcoholism, alcohol abuse; physical dependence
Addison disease (adrenal hypofunction), 143, 218–219, 220F, 247, 249, 399
  hypermagnesaemia, 255–256
adenine, 256F
adenine (A), 402, 406–407, 454F, 455F, 468B
adeno-associated virus (AAV), 414
adenocarcinomas, 495, 552T, 555
adenoma, 205, 209–210
adenomatous polyposis familial (FAP), 553
adenosine (A), 71–74
  see also deoxyribonucleic acid
adenosine deaminase (ADA), 132F
adenosine triphosphate (ATP), 89, 196–197, 198F
  angina, 438–439
  anthelmintic drugs, 93
  aspirin toxicity, 367, 370
  blood glucose regulation, 210
  chromosomes and inherited cancers, 522
  formation, 335–336
  RAD51-DNA complex, 458F
  RTKs, 197, 198F
  SDIS, 503
  see also energy
adenosine triphosphate (ATPase), 292B
adenovirus particles, 495
adenoviruses, 1F, 60, 69, 495
adenylate cyclase (AC), 197F, 211
adenylate cyclase transduction system, 39–40, 40F
ADH *see* alcohol dehydrogenase; antidiuretic hormone
adhesins, 36
adipokines, 528
ADP ribosylation, 39–40, 40F
adrenal carcinomas, 222
adrenal cortex, 216–224, 434M, 500
  aldosterone, 217–218
  cortisol, 217
  disorders, 218–223
  histological structure, 216F
adrenal gland, 436–437
adrenal insufficiency, Zellweger spectrum disorder, 502
adrenal medulla, 223–224B
adrenaline, 192, 193, 193F, 196
adrenergic β-blockers, 431, 449F
  see also β-blockers
adrenocorticotrophic hormone (ACTH), 200F, 201F, 204, 217F
  Addison disease, 218–219, 219F
  CAH, 221
  Cushing syndrome, 221
adrenoleukodystrophy (ALD), 499–500, 502
ADT *see* androgen deprivation therapy
adult Gaucher disease (GD1), 490
adult Refsum disease (ARD), 500
adult/maternal Hb (HbA), 387–388, 388F, 400B, 401B, 401F, 405B, 416C
advanced glycation end products (AGEs), 213, 573
aetiology, 2–3, 4F
affinity, 112
aflatoxin, 363B, 363F, 524T
aflatoxin B1, 524
AFP *see* α-fetoprotein
African trypanosomiasis, 502B
African trypanosomiasis/sleeping sickness, 43, 56, 85, 86T, 88
AFS *see* aqueous fluid sample
agammaglobulinaemias, 133, 134M
ageing, 295B, 360M
  age-related clinical conditions, 575–581
  causes, 571–575
  cells, 569–570
  epidemiology of cancer, 515F
  investigating/managing of clinical conditions, 582, 583T
  organs, 569, 570
  and SIDS, 129, 130T
  studying, 569M
  systems, 570–571
  tissues, 569, 570
agent dimercaptosuccinic acid (DMSA), 377
AGEs *see* advanced glycation end products
agglutination, 145F, 162F, 168F
AHG *see* antihuman globulin
AI *see* artificial intelligence
AIDS *see* acquired immunodeficiency syndrome
air pollution, 555
AJCC *see* American Joint Committee on Cancer
AL *see* amyloid lightchain
ALA *see* α-linolenic acid
alanine, 335–336, 335F
alanine transaminase (ALT), 305, 353B
alarmins, 125F
albendazole, 75M, 93F, 359
albumin, 112F, 246M, 253B, 253F, 311C
alcohol, 365, 435–436B, 449
  and cancer, 526, 528
  myocarditis/viral cardiomyopathy, 436
  see also ethanol
alcohol dehydrogenase (ADH), 369–370, 370F, 371
  and ethnicity, 370M
  ethylene glycol toxicity, 372, 372F
  methanol toxicity, 372
alcoholism, alcohol abuse, 298, 298B, 299, 353, 354C, 380C, 435
  see also chronic alcoholism; cirrhosis
ALD *see* adrenoleukodystrophy
aldehyde dehydrogenase (ALDH), 369–370, 370F, 372F
aldosterone, 217–218, 243, 247, 248, 434M, 500
  biosynthesis, 221, 222F
  Conn syndrome, 222
  metabolic acidosis, 269
alectinib, 545–546
α subunits, 101F
aliquots, 20C
alkaline phosphatase (ALP), 353B
alkalis, digestion, 326–327B
alkalosis, 221, 249, 251, 263, 268, 269T
alkylating agents, 182, 524T, 537, 538
all-*trans*-retinoic acid (ATRA), 471–472B
alleles, 453, 459
  see also chromosomes; mutations
allele-specific oligonucleotides (ASOs), 478–479
allergens, 148–149
allergic asthma, 148, 148M, 148T, 149–151
  see also asthma
allergic rhinitis *see* hay fever
allergies, 99F, 120, 151, 328, 448
  classification/effects, 148T
  definitions, 147
  food allergies, 151
  and intolerance, 334–335B
  type I hypersensitivities, 148
  type IV (delayed-type) hypersensitivies, 154
  see also hypersensitivities; immunological hypersensitivities
Allison, James, 544M
allogeneic HSCT, 185
allogenic gene therapy, 500
allografts, 174–175
allommunisation, 172
allopurinol, 49B, 256, 258F
allotopic expression, 499
ALLs *see* acute lymphoblastic leukaemia
all-*trans*-retinoic acid (ATRA), 471B
allylamines, 89, 90F
ALP *see* alkaline phosphatase
alpelisib, 517M, 545–546

αβ heterodimers, 139
α, 102F
α cells, 211
α chains, 395M, 401B
α genes, 401, 406–407
α globins, 401B, 406
α linolenic acid (ALA), 281–282
α particles, 529
α phytanoyl CoA hydroxylase (PAHX), 500
α polypeptides, 389F
α subunits, 139F
α subuntis, 387–388, 387F
α thalassaemias, 406–407, 408
  characteristics, 406T
α-1-antitrypsin, 272–273B
α-fetoprotein (AFP), 533T
α-linolenic acid (ALA), 527
α-tocopherol, 289–290, 289F
  see also vitamin E
ALT *see* alanine transaminase
alternative pathway LPS (complement), 104–106
aluminium, poisoning, 375, 376B
aluminium sulphate ($Al_2(SO_4)_3$), 376B
Alzheimer disease (AD), 5, 494–495, 520B, 575–579, 581M
amenorrhoea, 202, 204, 221–222, 230–231, 230F, 297
American Joint Committee on Cancer (AJCC), 536
American trypanosomiasis, 86T
  see also Chagas disease
Ames test, 526F
amines, 192
amino acids, 265, 283T, 454–455, 526
  absorption, 328, 329F
  blood clotting, 389F
  cachexia, 531–532B
  chemokines, 103F
  phenylketonuria, 464–465
  see also deoxyribonucleic acid; mutations; proteins; ribonucleic acid
aminoacyl-tRNA, 78
aminoglycosides, 78, 243
aminolevulinic acid (ALA) dehydratase, 376–377, 377F
5-aminosalicylic acid, 345F
amitriptyline, 324B, 435–436B
AML *see* acute myeloid leukaemia
ammonia, 335–336, 335F
ammonia (NH_3), 267, 340
amniocentesis, 136, 466, 477F, 489–490, 490F
amniocytes, 500
amodiaquine, 41–42B, 42F, 86
amoebiasis, 43, 85, 86T, 89
amoebic dysentery, 85, 86T
amoebicides, 86
amoxicillin, 56, 67, 337, 340–341
amphotericin B, 75M, 89, 90–91, 91F
ampicillin, 56
amplification, 104M, 389, 389M, 529
AMR *see* antimicrobial-resistant parasites
amyl nitrite (poppers), 395
amylase, 15, 320, 321F, 323, 327, 348
amylodoisis, 437, 437B
amyloid A protein, 439
amyloid and amyloidoses, 576–577B
amyloid lightchain (AL), 576B
AN *see* anorexia nervosa
anaemia, 299–300, 302, 311C, 343
  pernicious anaemias, 300, 323M, 340, 349–350, 399
anaemia of chronic disease (ACD), 398
anaemias, 160, 187, 305–306, 353, 353C, 354C, 385B, 387, 396–399, 401–402, 406F, 417C, 448, 506–508, 561, 564C
  carbon monoxide poisoning, 374
  classification, 396–397, 397T
  complement deficiencies, 140
  fetal, HDN, 170
  G6PD, 408–411
  haemolytic anaemias, 300, 410–411, 506–508
  see also sickle cell anaemia; thalassaemias
anaerobic glucose metabolism, 265
anaesthetics, 395
anagrelide, 415F
anakinra, 135B
analgesics, 243
analytes, 2
analytical sensitivity/specificity, 15, 16–17B
anaphase, 456F, 473
anaphylactic shock, 148T, 152B
anaphylatoxins, 106
anaplastic lymphoma kinase (ALK), 536, 545–546
ANAs *see* antinuclear antibodies
anastomoses, 424
anastrozole, 551–552, 551F, 551M
anaylsers, 12
androgen deprivation therapy (ADT), 558–559
androstenedione, 226–227
aneaemia, 349
anemias, 490
aneuploidy, 473–476, 473T
aneurysm, 445
Angelman syndrome, 470
angina, 435, 436B, 438–439, 449C
angiogenesis, 103, 514T, 525B, 527, 546–547
angiograms, 95C
angiography, 429–430, 439
angioplasty, 439
angiotensin, 392
angiotensin I/II, 217F
angiotensin-converting enzyme (ACE), 433–434, 434M, 449
anisocytosis, 398
anogenital warts (condyloma acuminata), 60
Anopheles mosquito, 41–42B, 42F, 48
  see also malaria
anorexia, 532B
anorexia nervosa (AN), 294, 296–297, 358B
  general management, 309
  investigation and laboratory testing, 304
anosmia (loss of smell), 230M, 316M, 500
anterior pituitary, 191F, 200–201, 201F
  CAH, 221
  cortisol, 217F
  growth hormone disorders, 204–206
  oral contraceptive pill, 229B
  reproductive hormones, 225–228
  thyroid hormone disorders, 206–207, 207F
anthelmintics, 92–94, 359
anthracyclines, 540
anthropometric tests, 303, 532B
  anorexia and bulimia, 304
  PEM, 303
anti-angiogenic drugs, 546M, 547
anti-D prophylaxis, 171
anti-immunoglobulin, 535
antiandrogen, 543T
antiandrogenic drugs, 230
antiapoptotic genes and their products, 519B, 527, 546
antiarrhythmic drugs, 431–432
antibacterial drugs, 76
  see also individual antibacterial drugs...
antibiotic therapy, 34, 35B, 39, 50C, 60
antibiotics
  acute myeloid (myeloblastic) leukaemia, 561
  biochemical and molecular biological techniques, 71F
  Chédiak–Higashi syndrome, 138
  controlling the spread of pathogens and parasites, 69
  drug resistance, 61–62B
  ear infections, 55–56
  eye infections, 55
  fat-soluble vitamins, 300–301, 301M
  infections of the urogenital system, 60, 61–62B
  inflammation, 99F
  penicillins, 80, 80M
  superantigens and toxic shock syndrome, 124B
  syphilis, 65
  treatment of infectious diseases, 75–94
  ulcers, 340–341
  see also immunoglobulins; *individual antibiotics...*

antibodies
ABO blood group, 164–165
apoptosis and cancer, 520B
blood groups/laboratory determination, 168
Duffy blood group system, 167
graft rejection, 176, 177B
infections of the respiratory system, 57
Kell blood group system, 166
Kidd blood group systems, 167
Lewis blood group system, 167
phage display libraries, 121B
serological HLA typing, 179–180
serological tests, 70–71
T cell depletion, 186
see also ABO blood group; immune system, immune response; immunoglobulins; monoclonal antibodies; Rhesus systems
antibody deficiencies, 130, 133–136
antibody–drug conjugates (ADCs), 545
anticancer drugs, 91, 534F, 537, 538F, 541F, 551F
hypomagnesaemia, 255
see also chemotherapy; *individual anticancer drugs...*
anticancer properties of curcumin, 527
anti-CCP *see* anti-cyclic citrullinated peptide
anti-CD2 MAbs, 185, 188C
anti-CD3, 131T
anti-CD3 MAbs, 185
anti-CD25 MAbs, 185
anticoagulant therapy, 435–436B
anticoagulants, 252, 283, 301M, 394B
anticonvulsant drugs, 299–300
anti-cyclic citrullinated peptide (anti-CCP), 146
antidepressants, 435–436B
antidiphteria toxin, 50C
antidiuretic hormone (ADH), 200F, 202F, 203–204, 217F, 243F, 246–247
hyponatraemia, 247
see also vasopressin
antierythrocyte antibodies, 162F, 169, 170, 171
antifungal drugs, 89–92
anti-GBM antibodies, 143–144
antigen–antibody reactions, 350–351
antigen-presenting cells (APCs), 110, 123F, 175, 185, 544, 553
immune checkpoint inhibitors, 544
antigens, 111
CAR-T, 544–545
Duffy blood group system, 167
graft rejection, 175
Kell blood group system, 166
Kidd blood group systems, 167
Lewis blood group system, 166–167
molecular biology HLA typing, 180F
prostate cancer, PSA, 533T, 557, 558, 563C
serological HLA typing, 179–180
superantigens, 41, 65–66, 124B
TAA, 544–545, 545F
tumour markers, 533T
see also ABO blood group; immune system, immune response; Rhesus systems
antiglobulin/Kleihauer testing, direct, 171
antiglomerular basement membrane disease, 143–144, 213
antihuman globulin (AHG), 171
anti-immunoglobulin, 131T
anti-inflammatories, 283, 539–540, 540F
Crohn disease, 345
see also nonsteroidal anti-inflammatory drugs; steroidal anti-inflammatory drugs
antilymphocyte globulins (ATG), 185
antimalarial drugs, 41–42, 42F, 395, 409
antimetabolites, 537
antimicrobial-resistant (AMR) parasites, 89
antimicrotubular agents, 539–540, 541F
antinausea drugs, 532B
antineoplastic agent, 405
antinuclear antibodies (ANAs), 146
antinuclear factors, 144
antiobiotics, and toxicity, 76M
antioxidants, 287, 289, 331, 527, 572–573, 581
antiphagocytic activity, 37–38
antiprotozoal drugs, 85–89
antisense oligonucleotides, 506, 506M
antistaphylococcals, 124B
antithrombic properties, 283
antithrombin III, 435B
antiviral drugs, 54, 54B, 69, 75–76, 138
see also individual antiviral drugs...
Antonine Plague, 30
ANUG *see* acute necrotizing ulcerative gingivitis
anus, 315, 316F
aorta, 424
aortic dissection, 445–446
aortic valve stenosis, 437–438
aortic valves, 430, 433F, 437–438
Apaf-1 *see* apoptotic protease-activating factors
APCs *see* antigen-presenting cells
apheresis, 132–133B, 173M, 405
API test strips, 70B
apical membrane copper transporter 1 (CTR1), 332–333
apical membrane transporter SR-B1, 332
apixaban, 394B, 394F
APL *see* acute promyelocytic leukaemia
apo *see* apolipoprotein-A through -E
apolipoprotein-A through -E, 440–441
familial hypercholesterolaemia, 441
apoptosis, 9F, 86, 103, 110, 132, 471B, 487, 495–496
and cancer, 516, 517, 519–520B, 527
chemotherapy, 540
targeted therapy, 546
apoptosome, 519–520B, 520F
apoptotic protease-activating factors (Apaf-1), 519B, 520F
appendix, 336F
aqueous fluid sample (AFS), 562–563B
arachidonic acid, 282F
ARD *see* adult Refsum disease
ARDS *see* acute respiratory distress syndrome
arginase, 293F
arginine, 39–40, 40F, 344B
arginine vasopressin, 243F
aromatase, 551–552, 551F, 551M
aromatase inhibitors, 543T
arrhythmias, 273B, 429, 430, 431–432, 435, 505
arsenic, 375, 378–379B, 415, 524T
poisoning, 377–378
artemisinin, 41–42, 42F, 86
arteries, 383, 424–425, 424F, 438
see also atherosclerosis
arterioles, 424F
arteriosclerosis, 438
arthralgia, 43, 146, 147, 405
arthritis, 575, 579
see also rheumatoid arthritis
arthropods, 48–50
artificial heart valves, 430, 433F
artificial intelligence (AI), 548B
aryl sulphatase, 489F
asbestos, 524T, 525B, 555
ascariasis, 47
Ascaris lumbricoides, 47F, 94F
ascites, 352–353, 353F, 534
ascorbic acid, 284F, 287
see also vitamin C
ASOs *see* allele-specific oligonucleotides
aspartate transaminase (AST), 305, 353B, 442, 443B
aspergilloma, 58
aspergillosis, 45, 90–91
Aspergillus flavus, 363B
Aspergillus fumigatus, 58
Aspergillus spp., 138
aspirin, 367–369, 367F, 379C, 415, 435–436B, 436F
overview of toxic effects, 367, 368F
poisoning
clinical features, 368
laboratory testing, 368–369
management, 369
toxic doses, 367

Association for the Advancement of Blood and Biotherapies (AABB), 173M
AST *see* aspartate transaminase
asthenia, 532B
asthma, 45, 109M, 120, 148, 148M, 148T, 149–151, 271, 273B, 275C
ataxia, 300, 490, 500
ataxia telangiectasia-mutated (ATM) kinase, 521, 536
ataxia-telangiectasia (*ATM*), 9F
atenolol, 431, 449F
atezolizumab, 544
ATG *see* antilymphocyte globulins
atheroma, 438
atherosclerosis, 300, 438, 439, 444
  cholesterol/lipoproteins, 439–440
athlete's foot, 125
ATM *see* ataxia telangiectasia-mutated kinase
ATM (ataxia-telangiectasia), 9F
atopy, 148M
atorvastatin, 359, 441–442
ATPases, 332–333
ATP-binding cassette (ABC) proteins, 441M, 483, 484F, 499–500
ATP-binding domains, 483
ATRA *see* all-*trans*-retinoic acid
atrial fibrillation, 431–432
atrioventricular (AV) node, 422, 423F, 428, 429, 432
atrium, 421F, 422, 429, 431–432
  see also heart
atrophic gastritis, 339
atypical chemokine receptor 1 (ACKR1), 167
Aurelius, Marcus, 30
auscultation, 422M
autoimmune disorders, 103M, 123, 129, 141–147, 200
  antiglomerular basement membrane disease, 143–144
  apoptosis and cancer, 520B
  diabetes mellitus/Addison disease, 143
  and endocrine glands, 142
  examples, 142T
  Graves' disease, 141, 142–143
  myasthenia gravis, 141, 144
  rheumatoid arthritis, 140, 144–146
  SLE, 144, 146–147
  thrombocytopenia, 415
  thryroiditis/Hoshimoto thyroiditis, 142, 143F
  type II hypersensitivities, 152
autoimmune gastritis, 340
autoimmune haemolytic anaemias, 140
autoimmune response, arthritis, 579
autologous HSCT, 185
automated analysers, 8F
automated blood group analysers, 168
automation, 12
autophagy, 519B
autosomal monosomy, 473
autosomes (nonsex chromosomes), 453
autumn crocus (*Colchicium autumnale*), 539–540, 540F
AV *see* atrioventricular node
avanafil, 233F
avelumab, 544
avidin, 299F
avidity, 112
AZA *see* azathioprine
azathioprine (AZA), 180–181, 181M, 182F
azidothymidine (AZT), 64B, 64F
azithromycin, 78

B

B cell activating factor (BAFF), 101–102
B cell lymphoma (BCL2) protein family, 518, 519B, 520B, 520F, 533T, 546
B cell survival factor, 101–102
B lymphocytes, 9, 9F, 102–103, 110F, 117, 123, 125, 384F
  activation of, humoral immunity, 120
  ALLs, 559–560
  antibody deficiencies, 134–135
  blood and blood products for transfusion, 161–162
  combined T and B lymphocyte deficiencies, 131
  distinguishing between B and T lymphocytes, 118–119B
  hairy cell leukaemia, 560M
  hyperimmunoglobulin syndromes, 135B
  immunosuppressants/immunosuppression, 185
  marker proteins, 118T
  production of specific immune responses, 117–118
  SCID, 130, 131–133
  viruses and cancer, 529
B7 homologue 1 (B7-H1) *see* programmed death-ligand 1
babies *see* neonates
Bacillus Calmette–Guérin (BCG) vaccine, 153, 543
Bacillus polymyxa, 82
bacteria, 2F, 24
  antibody deficiencies, 133–134
  cardiac failure, 433
  chemical carcinogens, 526F
  defensive virulence factors, 35–36
  dietary malabsorption, 343
  drug resistance, 61B
  endocarditis, 430
  endoplasmic reticulum, disorders of, 487
  food poisoning, 360M
  gastritis, 339–340, 339F
  helper T lymphocyte activation, 123
  immunodeficiency diseases, 129
  immunotherapy, 543
  infections of the skin, 54–55
  lactose intolerance, 342
  microbial virulence factors, 34–41
  neutrophils, 108
  offensive virulence factors, 36–41
  P-450 systems, 363B
  penicillins, 80, 80M, 152
  plague, 34–35B
  platinum-based drugs, 541
  rheumatic fever, 432B
  sickle cell anaemia, 404, 405B
  stomach acid, 323
  thrombocytopenia, 415
  topoisomerase inhibitors, 540
  vitamin K deficiencies, 300–301, 301M
  see also individual bacterial species... infections, infectious diseases; mycobacteria
bactericidal drugs, 76
bacteriophages, 84–85B
bacteriostatic agents, 76
Bacteroides fragilis, 35–36, 36F
bad air *see* malaria
bad air disease *see* malaria
bad breath, 337
'bad' cholesterol, 441
BAFF *see* B cell activating factor
balanced diet, 280
Baltimore classes (BCs) of viruses, 28T
Baltimore, David, 28, 28M
band ligation, 348
barbiturates, 373, 435–436B
barbituric acid, 373F
barium sulphate, 338
Barnard, Christiaan, 173
Bart hydrops, 406T, 407–408
basal cell carcinoma, 529
basal lamina, 503F
basophilic stippling, 380C
basophils, 108, 109, 384F, 386, 562
  possible problems, 109M
  see also mast cells; megaloblasts
BAt *see* brown adipose tissue
BBMR *see* British Bone Marrow Registry
BCG *see* Bacillus Calmette–Guérin; Bacillus Calmette–Guérin vaccine
BCL11A, 388F, 389M, 405
BCR gene mutations, 522, 562
  see also Philadelphia chromosome
Becker muscular dystrophy (BMD), 504–506, 504T
Becker, Peter, 505

bee venom, 152F, 154C
belinostat, 547
benign pseudohypertrophic muscular dystrophy *see* Becker muscular dystrophy
benzene, 415, 524T, 561
benzimidazole anthelmintics, 93
benzo[a]pyrene, 363B, 363F, 524T, 525, 555
benzocaine, 395
benzylamines, 89, 90F
beriberi, 298, 298B
β agonists, 151
β antithrombin III, 394F
β-blockers, 324B, 431, 431F, 434–435, 437, 439, 443–446, 449F
β carotene, 287–288, 288F, 340
β cells, 212, 216
β chains, 395M, 401B, 403
β globins, 388F, 401, 402–403, 405, 406, 407
β globulin genes, 405, 408
β glucans, 125
β glucocerebrosidase, 490, 491F, 508C
β haemolysins, 38, 39F
β hydroxybutyrate, 210, 265F
β lactamases, 80
β lactams, 80F
β microglobulin (β_2M), 177
β oxidation, 285
β particles, 529
β polypeptides, 177–178, 389F
β subunits, 101F, 139F, 387–388, 387F
β thalassaemias, 406, 407, 408
β tubulin, 93
β tubulin toxin, 360
β2 integrin, 139
β_2 integrin (CD18), 139
β_2 microglobulin (β_2M), 122
β-amyloid peptides (Aβ), 575, 576F, 577–578
β-catenin, 553
β,γ complex form of MgATP, 255F
bevacizumab, 546M, 547, 553
bile, 318, 327–328, 343, 353, 385
bile canaliculi, 350
bile duct, 317F
bile duct disorders, 349–353
bile salts, 371, 439–440
bilharzia *see* schistosomiasis
biliary cirrhosis, 353, 494–495
bilirubin, 10–11B, 10F, 170, 317–318, 318F, 349–350, 385, 507–508
bilirubinuria, 351
binding, 29F
biochemical and molecular biological techniques, 71
biological variation, 17
biomarkers, 433, 442–443, 443B
  tumour markers, 531–536, 545
bioprosthetic valves, 175
biopsy, 13–14
biopsy, tumour cells, 534–535
biosynthesis
  alkylating agents, 538
  antiprotozoal drugs, 87–89
  bacterial cell wall, inhibitors, 79–81
  cortisol/aldosterone, 221, 222F
  familial hypercholesterolaemia, 441–442, 442F
  leptin biosynthesis, 532B
  leptin, cachexia, 532B
  mitochondrial matrix, 495–496
  topoisomerase inhibitors, 540
biotin, 286, 287F
  see also vitamin B_7
bird flu, 31–32
Bismarck brown, 535
bismuth chelate, 340–341
bisoprolol, 431, 449F
bisphosphate pamidronate, 252F
2,3 bisphosphoglycerate (BPG), 396
Black Death, 31, 35, 35B
bladder cancer
  cigarettes and smoking, 525–526B
  immunotherapy, 543
  platinum-based drugs, 541
  topoisomerase inhibitors, 540
blast cells, 559F
blastomas, 514
bleeding time, 391
blood
  acute myeloid (myeloblastic) leukaemia, 561
  collection, 14–15, 14F
  cytokines, 103
  glucose levels, 210–216, 210F
  glucose meters, 215F
  nonspecific defences, 100
  prostate-specific antigen, 557, 558, 563C
  tumour cells, 534–535
  vessels, angiogenesis, 103, 514T, 525B, 527, 546–547
  see also haemoglobin; immune system, immune response
blood banks, 159F, 160
blood clotting, 290, 317, 389–395, 435, 443–444
  angina, 439
  atherosclerosis/arteriosclerosis, 438
  clinical aspects, 411–415
  clot dissolution, 395
  clotting factors, 389–391
  clotting pathways, 391–393
  clotting time, 392
  factor classification, 390–391
  graft rejection, 176
  pathways, 391–393
  prevention, 394
  thrombolysis, 439, 439M
  see also individual factors...
blood count, 70, 147T, 311C, 343, 398, 433, 560M, 562, 564C
blood disorders, 383–419
  anaemias, 387, 396–399
  blood composition, 383T
  clotting, 389–395
  erythrocytes, 384–385, 384F
  G6PD, 408–411
  haemoglobin, 387–388
  haemoglobinopathies, 400–408
  leukocytes, 384F, 386
  methaemoglobinaemia, 395–396
  platelets, 383T, 384F, 386, 387F
blood gas analysers, 274F
blood groups, 162–165, 166–167, 188C
  ABO blood group, 152, 153F, 162–165, 168, 170–172, 186–187, 188C
  genotype/phenotype, 458
  laboratory tests, 168
  type II hypersensitivies, 152
  see also ABO blood group
blood pressure, 235C, 424, 426–428, 450C
  angiotensin-converting enzyme, 434M
  cardiac failure, 433
  clinical problems, 447–449
  measuring, 426–428
blood tests, acute hepatitis, 350–351
blood transfusion, 152
  see also transfusion and transplantation
blood transfusions, 448
  hepatitis, 352
  iron metabolism, 385B
blood vessels, 424–426
  see also arteries; veins
blood–brain barrier, 56F, 86–88
  barbiturates, 373
  mitochondrial disorders, 499
  pineal gland, 194B
blood-clotting factors, 290, 317
BMD *see* Becker muscular dystrophy
BMI *see* body mass index
BMT *see* bone marrow transplant
body mass index (BMI), 294, 295B, 303F
bolus, 322
bone, 249, 290–291, 293
  hypervitaminoses, 301
  mineral/trace elements, nutritional deficiencies, 301, 302, 302B
  osteomalacia, 300
  resorption, 249
  turnover, 251F
  see also osteoporosis; teeth; vitamin D
bone marrow, 9F, 416C
  acute myeloid (myeloblastic) leukaemia, 561

acute promyelocytic leukaemia, 472B
anaemias, 397, 398, 399F
blood disorders, 384F
chemotherapy, 537
chromosomes and inherited cancers, 522
chronic myelogenous leukaemia, 562
Gaucher disease, 490F
hereditary elliptocytosis/spherocytosis, 506
kidney disorders, 239–240
thalassaemias, 406
thrombocytopenia, 414–415
tumour cells, 534–535
water-soluble vitamins, investigation, 305–306
bone marrow transplant (BMT), 132, 132–133B, 137, 185, 186–187, 491
post-transplant support, 187
Bordetella pertussis, 37–38, 57, 58F, 67–68
Borrelia burgdorferi, 48F, 66, 73
Borrelia spp., 66–67
bortezomib, 546
bosutinib, 562
botulinum toxin, 39F, 360
botulism, 39
bowel cancer, 11–12, 12–13M, 247, 347–348
bowel cancer testing kits, 533F, 533M
Bowman's capsule, 242
BPG *see* 2,3 bisphosphoglycerate
brachytherapy, 542, 558
bradycardia, 430
bradykinin, 392
brain
Alzheimer disease, 575–579
angiotensin-converting enzyme, 434M
atrial fibrillation and flutter, 431
cancer, 324B
carbon monoxide poisoning, 374
Creutzfeldt–Jakob disease (CJD), 26F, 160T, 462–463B, 462T, 463T
cryptococcal meningitis, 90–91
ethanol toxicity, 369
ethylene glycol toxicity, 372
kernicterus, 170, 349
phenylketonuria, 464
pineal gland, 194–195B, 194F
proton beam therapy, 542
spongiform encephalopathies, 2M, 26, 160T, 462–463B, 462–463T
stroke, 446F
Tay–Sachs disease, 489–490, 489T, 490F
thrombocytopenia, 415
thrombocytosis, 415
Wernicke–Korsakoff syndrome, 298B
Zellweger spectrum disorder, 502
see also central nervous system; *individual brain structures…*; stroke
Branca, Antonio, 174B, 174F
breast cancer, 121B, 495, 514, 515, 515T, 517M, 531, 533–534, 533T
and alcohol, 528
antimicrotubular agents, 540
curcumin and its anticancer properties, 527
diagnosis, staging and treatment, 549–552
hormone therapy, 542–543
Li–Fraumeni syndrome, 522B
and obesity, 528
risk factors, 549
signs and symptoms, 549
targeted therapy, 546
TNM system, 550T
topoisomerase inhibitors, 540
breast cancer type 1 susceptibility protein (BRCA1), 518, 521, 536, 546, 549, 553, 557
breastfeeding, 202–203, 202F, 203T
see also prolactin
Brenner, Sydney, 520B
British Bone Marrow Registry (BBMR), 185M
broad bean plant (*Vicia fava*), 409, 411B, 411F
broad-spectrum antibiotics, 76
bromocriptine, 203F, 206
bronchiectasis, 134
bronchitis, 273B, 275C, 525B, 555
bronchoscopy, 534F, 555
brown adipose tissue (BAT), 532, 532B
Bruckner test, 562B, 562F
brush border, 325F
Bruton tyrosine kinase (BTK), 134–135, 134F, 136
BTK *see* Bruton tyrosine kinase
bubonic plague, 35B
Buckley–Job syndrome, 135B
buffering, 265–267, 268F
buffers, 213
bulimia nervosa, 296–297, 324B, 358B
general management, 309
investigation and laboratory testing, 304
bundles of His, 422
Burkholderia cepacia, 138
Burkitt lymphoma, 529
butenafine, 89, 90F
'butterfly' rash, 146F
butyrate, 404–405, 404F
butyric acids, 270F
bypass surgery, 308–309, 433, 439, 444F

C

C proteins *see* complement
C1 inhibitor (C1INH), 140, 140B
C3b opsonin, 139
Ca^{2+}-ATPase (Ca^{+} pump), 332
cabazitaxel, 540
cachectic, 298B
cachetin, 531–532B, 531F
cachexia, 107B, 107F, 294, 531, 531–532B
CAD *see* caspase-activated DNase
cadmium, 333, 375, 379
calcineurin, 183–184, 183F, 184F
calcitonin, 249, 533T
calcitriol, 239–240, 249, 289
calcium (Ca^{2+}), 15, 82, 93, 139F, 160F, 183–184, 183F, 196, 197F, 270, 290–291, 304–305, 305F, 307, 332–333, 332F, 333F, 441F
botulinum toxin, 39F
channel antagonists, 435, 437, 439
clinical aspects of clotting, 411–412
clotting pathways, 392–393, 392F
endoplasmic reticulum, disorders of, 487
ethylene glycol toxicity, 372
haemophilia A, 413F
homeostasis disorders, 249–253
mitochondrial disorders, 495–496
see also blood clotting
callipers, 303F
calmodulin (CAM), 183–184, 184F
calorie restriction, 581–582
calprotectin, 345F
CAM *see* calmodulin
Camelford, Cornwall, 376B
cAMP *see* cyclic adenosine monophosphate
cAMP-regulated Cl^{-} channels, 483, 484F
Campylobacter jejuni, 60F
CaN *see* cytosolic phosphatase calcineurin
cancer, 30, 73, 126C, 513–567
acute promyelocytic leukaemia, 471–472B
ageing, 569, 575, 579–580
anaemias, 398
and apoptosis, 517, 519–520B
carcinomas, 222, 247, 495, 514, 529, 533
causes, 521–531
cigarettes and smoking, 525–526B
classification, 514
COPD, 273B
cytokine release syndrome, 103M
cytokines, clinical features, 154B
deep vein thrombosis, 447
definitions, 513
disorders of the exocrine pancreas, 348
disorders of vitamin nutrition, 299
endoplasmic reticulum, disorders of, 487
epidemiology, 515
fragile sites, 468
Gaucher disease, 490
general diagnosis, 531–536
general treatment, 536–548
chemotherapy, 536, 537–541
gene therapy, 536, 547–548

cancer (*continued*)
hormone therapy, 536, 542–543
immunotherapy, 542, 543–545
radiation therapy, 536, 541–542
targeted therapy, 545–547
genetics and DNA, 454
Golgi apparatus, disorders of, 488
haemorrhoids, 347–348
hiatus hernias, 338
HLA typing, 179
hypomagnesaemia, 255
immunodeficiency diseases, 129
immunosuppression, 181
infections of the skin, 54
internal/external signals, 513F
intraheptic gallstone obstruction, 350
lactose intolerance, 342
molecular biology, 515–521
monoclonal antibodies, 121B
mouth disorders, 337
myeloma, 9F, 126C, 490, 543
notifiable diseases, 6
nucleocytoplasmic transport, 495
PCR, 74
PEM, 294
pheochromocytomas, 223–224B
plasma membrane-related disorders, 483
screening tests, 12–13M
specific types, 548–563
breast cancer, 549–552
chronic myelogenous leukaemia, 562
colorectal cancer, 552–554
leukaemias, 559–563
lung cancer, 555–556
prostate cancer, 556–559
storage of transplant material, 187F
thrombocytopenia, 415
triazoles, 91
vitamin E deficiency, 300
see also chemotherapy; leukaemias; myeloma; tumours
cancer antigens 19-9 and 125 (CA19-9, CA125), 533
cancer centres, 542F
Cancer Gene Census (CGC), 516
cancer genes, 517
Candida albicans, 45–46, 58, 59F, 60, 125, 129, 181
hyperimmunoglobulins syndromes, 135B
mouth disorders, 337
Candida auris, 45–46
Candida infections, 90
Candida spp., 94
candidiasis, 43, 45–46, 90–91
Candida guilliermondii, 45F
canines, 321B
cankers, 337
capillaries, 383, 424F, 426
capillary zone electrophoresis (CE), 407–408
capsids, 27F, 57F, 173F, 495
hepatitis, 351, 352F
capsomeres, 27
capsules, 35–36, 36F
captopril, 434F, 434M
carbamazepine, 299–300
carbimazole, 210F
carbohydrases, 40
carbohydrates, 105F, 192, 214–215, 280F, 281, 343, 345
absorption, 328
dietary malabsorption, 343, 345
mouth, digestion, 320, 321F
carbon dioxide (CO_2), 340, 341F, 422
hypotension, 448
pressure (PCO_2), 213, 266B, 268–269, 270, 271–272
blood gas analysers, 274F
mixed acid–base disorders, 272
pressure PCO_2, 367
production/transport, 263–265, 263F
see also hydrogen carbonate
carbon monoxide (CO), 373–374
carbonic acid, 265–267
carbonic anhydrase, 263–264, 264F, 265M
carboplatin, 541F, 563B
carboxyhaemoglobin, 374F
carboxyl terminal consensus sequence SKL (-Ser-Lys-Leu-COO^-), 499
carbuncles, 54–55
carcinoembryonic antigen (CEA), 533, 533T, 535, 553
carcinogenesis, 363B, 363F
carcinogen-inducing tumours, 524F
carcinogens, 514, 555
carcinoid tumours, 552T
carcinoma *in situ* (CIS), 536
carcinomas, 12F, 222, 247, 495, 514, 529, 530F, 533, 552T, 555
cardiac arrest, 269, 297
cardiac biomarkers, 433
cardiac catheterisation, 429–430, 434
cardiac catheters, 439
cardiac failure, 396, 432–434, 449
cardiac ischaemia, 429, 438–439
cardiac muscle complex, 443, 444F
cardiac oxygen and angina, 438–439
cardiomyopathies, 435–437, 505
cardiovascular disorders, 300, 421–452, 487, 575
abnormal heart rhythms, 430–432
arteries and veins, 445–447
atherosclerosis/arteriosclerosis, 438–442
blood pressure, 426–428
blood vessels, 424–426
cardiac failure, 432–434
cardiomyopathies, 435–437
endocarditis, 430
heart and circulatory system, 421F, 422–423
heart tumours, 438
heart valve disorders, 437–438
investigation, in health and disease, 428–430
ischaemic heart disease, 434–435
myocardial infarctions, 429, 432M, 433, 436B, 439, 442–444, 447–448, 449, 450C
pericardial disease, 445
cardioversion, 431–432
carditis, 432B
carotenoids, 527F, 557
carriers, 24M
CAR-T *see* chimeric antigen receptor T cell therapy
carvedilol, 449
Cas nuclease, 548
cascular endothelium, 403
caspase-activated DNase (CAD), 519–520B
caspases, 102–103, 519–520B, 520F, 527
catalases, 40, 374–375, 375F, 573F
Catalogue of Somatic Mutations in Cancer (COSMIC), 516
catalytic activity *see* enzymes
cataracts, 4, 575, 580
see also diabetic cataracts
catecholamine hormones, 223–224B, 287
catecholamines, 464
catechol-*O*-methyltransferase (COMT), 223B
cation (Ca^{2+}) channels, 233B
Cayler cadiofacial syndrome, 136M
CBC *see* complete blood count
CCK *see* cholecystokinin
CD3 proteins, 119B, 119F, 131B, 161–162
CD3-TCR complex, 185F
CD4+ and CD8+ cells, 62B, 141B, 142, 153–154
CD4 proteins, 28, 123, 134, 161–162
CD8 proteins, 122
CD15/sialyl Lewis X tetrasaccharide/glycan, 139F
CD18 proteins, 139, 155C
CD33 proteins, 561
CD34+ proteins, 187F
CD40 gene mutations, 135B
CD152 *see* cytotoxic T lymphocyte antigen 4
CD223 *see* lymphocyte activation gene-3
CD274 *see* programmed death-ligand 1
CD279 *see* programmed cell death protein-1
CDC *see* Centers for Disease Control and Prevention
CE *see* capillary zone electrophoresis
CEA *see* carcinoembryonic antigen

cefaclor, 99F
cefixime, 56
cell adhesion molecules, 122, 138–139, 139M
cell cycle, 455, 455F, 513F
  acute myeloid (myeloblastic) leukaemia, 561
  apoptosis and cancer, 519–520B
  arrest, 9F
  chemotherapy, 538, 539F, 540
  definitions, 455
  diet and cancer incidence, 527
  epigenetic regulators, targeted therapy, 547
  and malignancy, 456M
  overview, 455F
  retinoblastoma susceptibility gene, 518, 520
  tumour protein 53 gene, 518
  see also meiosis; mitosis
cell death, 365F
  see also apoptosis
cell differentiation, 102, 123, 384F, 518, 519T, 527
cell division, 455F
cell migration, 518, 519T
cell proliferation, 102, 123, 545–546
  see also cancer
cell-mediated immune response, 9, 100, 111, 118M, 123, 124, 125, 185
  antibody deficiencies, 134
  autoimmune disorders, 141, 144
  checkpoint inhibitors, 544
  classical WAS, 137
  DiGeorge anomaly, 136
  gastritis, 340
  graft rejection, 176–177
  immunological hypersensitivities, 153
  immunosuppression, 185
  rheumatoid arthritis, 144
cell-related disorders, 483
cell-to-cell communication, 518, 519T
cementum, 320–321B, 321F
cemiplimab, 544
Centers for Disease Control and Prevention (CDC), 56, 59, 62B, 272B
  HIV and AIDS, 62B
central nervous system (CNS), 36, 39, 193
  adrenoleukodystrophy, 499–500
  barbiturates, 373
  cryptococcosis, 46
  deep mycoses, 125
  ethanol toxicity, 371
  ethylene glycol toxicity, 372
  general treatment of cancers, 536–537
    chemotherapy, 537
  Hartnup disease, 344B
  hypercalcaemia, 251
  hypoglycaemia, 215
  immune response to parasites, 125
  infections of, 55, 56
  infections of the urogenital system, 65
  metabolic acidosis, 269
  methaemoglobinaemia, 395–396
  proton beam therapy, 542
  trypanomiases, 43
  trypanosome infections, 86
  Zellweger spectrum disorder, 502
  see also brain
centromeres, 456, 465F, 472
centrosomes, 456
Cephaelis ipecacuanha, 86
cephalosporins, 80
cerebral beriberi, 298B
cerebral cortex, 489–490
cerebrospinal fluid (CSF), 43
  pineal gland, 195B
cerebrovascular accident (CVA), 446
  see also stroke
ceruloplasmin, 333F
cervical cancer, 11–12, 12–13M, 62, 515, 530F, 534–535, 547
cervical screening programmes, 531M
cestodes, 46, 92F, 92T
cetuximab, 546, 553
CF *see* cystic fibrosis
CFTR *see* cystic fibrosis transmembrane conductance regulator
CFTR gene mutations, 484–486
  in mice, 486M
CGC *see* Cancer Gene Census
CGD *see* chronic granulomatous disease
cGMP *see* cyclic guanosine monophosphate
Chagas, Carlos, 49B
Chagas disease, 25–26, 43, 48, 49–50B, 85, 86, 86T, 87F, 502B
Chain, Ernst, 80M
chancre, 65
CHD *see* coronary heart disease
checkpoint kinase 2 (CHK2), 521
Chédiak–Higashi syndrome (CHS), 137–138
CHEK2/CHK2 gene mutations, 521
chelation, 302F
chelation therapy, 408
chelators, chelating agents, 172F, 252, 302F, 375–377, 404, 411
chemical carcinogens, 524–526
chemokine receptor CCR5, 103F
chemokines, 102T, 103, 103F, 167, 525B
chemotactic factor for eosinophils (ECF-A), 109T
chemotactic factor for neutrophils (NCF-A), 109T
chemotaxis, 103
chemotherapy, 45, 337, 531B, 536, 537–541
  acute myeloid (myeloblastic) leukaemia, 561
  alkylating agents, 538
  ALLs, 560
  AML, 561
  antimicrotubular agents, 539–540, 541F
  breast cancer, 551–552
  chronic myelogenous leukaemia, 562
  CLL, 560
  CML, 562
  colorectal cancer, 553
  hairy cell leukaemia, 560M
  lung cancer, 556
  platinum-based drugs, 541
  prostate cancer, 559
  purine and pyrimidine analogues, 538–539
  retinoblastoma, 563B
  topoisomerase inhibitors, 540
chenodiol, 350F
Chernobyl accident, 528
chiasmata, 458
chicken pox, 29F, 54, 415
childhood ALD, 500
childhood epiglottits, 56M
chimeric antibodies, 121B
chimeric antigen receptor T cell therapy (CAR-T), 544–545
chitin, 125
CHK2 *see* checkpoint kinase 2
Chlamydia, 36–37
chlamydia, 55
Chlamydia pneumoniae, 36–37, 58
Chlamydia spp., 36–37, 487
Chlamydia trachomatis, 55, 56F, 73–74
chloramphenicol, 76M, 79F, 415
chloride (Cl^-), 39–40, 93, 290, 320, 320F, 326B, 346B, 483, 485–486
chloride shift, 264–265, 264F
chlorpropamide, 204F
cholecalciferol, 289F
cholecystokinin (CCK), 336
cholelithiasis, 294, 295T, 350, 351B
cholera, 39–40, 40F, 85B, 485B
cholestasis, 349–350, 351B
cholesterol, 17F, 37–38, 192, 214–215, 289F, 330F, 332, 335–336, 349–350, 438, 439–442, 449C, 583C
cholesteryl, 439–441
choriocarcinoma, 540
chorionic villus cells, 500
chorionic villus sampling (CVS), 136, 466, 477F
chorioretinitis, 55
Christie Hospital, Manchester, 542F
Christmas factor (factor IX), 392
chromatids, 456–458, 457F

chromatin, 456, 519–520B
see also deoxyribonucleic acid
chromium, 291T, 301, 307
chromosomal banding, 466–467
chromosomal loops, 470–471
chromosomal mutations, 469–477
chromosomes, 8F
9 and 22, translocation, 522, 523F
ageing, programmed ageing, 574–575, 575F
complement of a normal human female, 466F
complement of a normal human male, 453F
definitions, 453
genomes, 459, 459M
haemophilia A, 413
and the human karyotype, 465–469
humoral immunity, activation of T lymphocytes, 122, 178F
and inherited cancers, 522–523
long chromosome 9, 522F
nuclei and nucleocytoplasmic transport, 491–492, 495
progeria, 581–582B
scanning electron micrograph, 453F
SCID, 132
sex-linked genetic diseases, 460–461
telomeres, telomerase reverse transcriptase, 516B
variations in chromosome numbers, 473–477
see also meiosis; mitosis; mutations
chronic alcoholism, 254, 298, 298B, 365, 369–371
chronic bronchitis, 525B
chronic diseases, 3
chronic gastritis, 339
chronic graft rejection, 175, 177
chronic granulomatous disease (CGD), 138–139
chronic heart failure, 433–434, 438
chronic infection, 107B
chronic kidney disease (CKD), 243, 244, 250, 255–256
chronic liver disease, 352
chronic lymphocytic leukaemia (CLL), 9F, 559, 560
chronic myelogenous leukaemia (CML), 495, 522F, 523F, 543, 559, 562
chronic obstructive pulmonary disease (COPD), 149, 150B, 272–273B
chronic pancreatitis, 348
chronic renal disease, 203–204
chronic ulcers, 340
CHS *see* Chédiak–Higashi syndrome
chylo, 440T
chylomicrons, 440–441
chyme, 323, 328
chymotrypsin, 326–327
cigarettes *see* smoking
cilgavimab, 121B
ciprofloxacin, 76
circulating tumour cells (CTCs), 536
cirrhosis, 350, 352–353, 371, 371F, 415
cisplatin (*cis*-diamminedichloroplatinum), 541F
cisplatinum, 255
cisternae, 487, 488F
citrate, 411
CJD *see* Creutzfeldt–Jakob disease
CK *see* creatine kinase
CKD *see* chronic kidney disease
CKMB *see* creatine kinase myocardial band
CKMM biomarker, 443B
clarithromycin, 78F, 340–341
class I/II HLA proteins, 177–178, 179
class I/II MHC proteins (MHC I/II), 175, 177
classical pathway for complement, 104–106
clindamycin, 78, 124B
clinical audits, 18–19
clinical biochemistry, 8
clinical blood grouping/typing, 168
clinical genetics, 8
clinical laboratories, 8–10
clinical malabsorption, 531B
see also malabsorption
clinical manifestations, 3–4
clinical specificity/sensitivity, 15, 16–17B
clinical tests for cancer, 532–536
CLL *see* chronic lymphocytic leukaemia
clofazimine, 76, 359
clofibrate, 435–436B
clonal selection, 118F
clonidine, 206F
cloning, monoclonal antibodies, 120–121B
Clostridium botulinum, 39, 360M
Clostridium difficile, 60, 81, 82
Clostridium perfringens, 38F, 40, 94F
clotting *see* blood clotting
clustered regularly interspaced short palindromic repeats (CRISPR), 545, 547–548
CMD *see* congenital muscular dystrophy
CML *see* chronic myelogenous leukaemia
CMV *see* cytomegalovirus
CO_2 production/transport, 263–265, 263F
coagulation disorders, 417C
see also anticoagulants; blood clotting
coagulation factors, 249
CoASH coenzyme A, 135B, 135F, 441–442, 442F, 577F
coatomer protein complex, α subunit (COPA), 488, 489T
cobalamin, 286–287
see also vitamin B_{12}
cobalt, 291T, 333
Coccidioides immitis, 94
coccidioidomycosis, 90–91
cochicine, 540F
cockroaches (*Periplaneta americana*), 23F, 48
codeine, 358F, 358M
codons, 400M, 403M
see also deoxyribonucleic acid
coeliac disease, 341, 342–343
dietary malabsorption, 343
coenzyme A, 286, 287F
coenzymes, 87F, 284, 285F, 359F, 362F, 387F, 408M
breast cancer treatments, 551F
CoASH coenzyme A, 135B, 135F, 441–442, 442F, 577F
ethanol toxicity, 371
FAD, 87F, 223B, 223F, 256F, 285F, 304–305, 305F, 362F, 387F, 409, 410F
familial hypercholesterolaemia, 441–442, 442F
G6PD, 409, 410F
macrocytic anaemia, 399
mitochondrial disorders, 499
NADH, 370, 372, 408M, 495–496, 497F
$NADP^+$, 286F, 305, 408–409, 409F, 411B
see also nicotinamide adenine dinucleotide phosphate
cofactor proteins, 390–391
coiled-coils, 488M, 503F
colchicine, 258, 466, 539–540, 540F
cold sores, 54B, 337
Coley's toxin, 543
colistin, 82
colitis, 60
collagens, 525B
collecting ducts, 240, 241F
collection, transport and storing of specimens, 13–14
colon, 345
colour blindness, 461T
colorectal cancer (CRC), 478, 528, 552–554
anti-angiogenic drugs, 547
chemotherapy, 540
cigarettes and smoking, 525–526B
curcumin and its anticancer properties, 527
and diet, 526–527
lactose intolerance, 342
risk factors, 552–554
signs, symptoms and diagnosis, 553
staging and treatment, 553, 554F, 554T
TNM system, 554T
combination therapy, 41–42B, 64B, 75, 75M, 76, 83, 86–88, 180–181

acquired mitochondrial DNA defects, 498B
breast cancer, staging and treatment, 550
colorectal cancer, staging and treatment, 553
definitions, 75M
platinum-based drugs, 541
common (third) final clotting pathways, 391F
common variable immunodeficiency (CVID), 134
communicable diseases, 23, 24M
compensation, 268, 269
complement, 103–107, 140, 175–176B, 389M
activation, 104–105
amplification, 104M
apoptosis and cancer, 519B
and blood transfusions, 169–170
opsonisation and inflammation, 106
serological HLA typing, 179–180
complementary determining regions (CDRs), 121B
complete blood count (CBC), 398
compulsive abuse or psychic craving, 373
computed tomography (CT), 7, 8F, 70, 429F, 446
AI in oncology, 548B
ischaemic heart disease, 434
lung cancer, 555
oncology, artificial intelligence, 548B
peritonitis, 320B
pineal gland, 195B
pituitary function, 202
prostate cancer, 557
radiation therapy, 541
stroke, 446
thalassaemias, 417C
tumour detection/localisation, 533
Wernicke–Korsakoff syndrome, 298B
COMT *see* catechol-*O*-methyltransferase
ConA *see* concanavalin A
concanavalin A (ConA), 131T
condensation, 198–199, 199F
conformational radiotherapy (CRT), 558
congenital adrenal hyperplasia (CAH), 218, 221–222, 230
congenital alactasia, 341–342
congenital diabetes mellitus, 203–204
congenital diseases, 5
congenital hyperthyroidism, 208
congenital muscular dystrophy (CMD), 504T
congestive heart failure, 505
conjugate vaccines, 68–69
conjugated bilirubin, 317–318
conjunctivitis (pink eye), 55
Conn syndrome, 218, 222, 248, 249, 270
conotruncal anomaly face syndrome, 136M
contact allergies, 154
contagious diseases, 23, 24M
contraception, 229B, 447
control samples, 18
Cooley, Thomas, 406
Coombs, Robin, 147, 148T
COPA *see* coatomer protein complex, α subunit
COPD *see* chronic obstructive pulmonary disease
copper (Cu^{+}), 160F, 291T, 292B, 301, 307, 332–333, 333F, 413F
corneal arcuses, 441, 450C
coronary heart disease, 205, 432, 435, 439, 440F, 441
see also ischaemic heart disease
coronary heart disease (CHD), 283, 294, 569
coronary infarction, 431
see also myocardial infarctions
coronary thrombosis, 411, 439, 525B
coronavirus, 28T, 32–33, 57
see also COVID-19
corpora, 233B
corticosteroids, 181T, 182F, 188C, 415, 437B, 506, 532B
immunosuppression, 180–182
rheumatoid arthritis, 146
corticotrophin-releasing factor (CRF), 217F, 532B
corticotrophin-releasing hormone (CRH), 201F
cortisol, 17, 192, 193, 217, 221, 222F, 500
Corynebacterium diphtheriae, 40, 50C, 57, 58F
COSMIC *see* Catalogue of Somatic Mutations in Cancer
COVID-19, 33F, 69, 74, 74F, 103M, 296, 297F
immunoassays, 116B, 116F
monoclonal antibodies, 121B
Coxsackie virus, 53F, 56, 212, 436
C-peptide, 211F, 216
cranial diabetes insipidus, 203–204
CRC *see* colorectal cancer
C-reactive protein (CRP), 70, 107B, 432B, 439
creatine kinase (CK), 442, 443B, 450C, 505F
creatine kinase myocardial band (CKMB), 443B, 443F
creatine phosphate, 438–439
creatinine, 12F, 243–244
cretinism, 208, 301, 307
Creutzfeldt–Jakob disease (CJD), 462–463B
see also variant Creutzfeldt–Jakob disease
CRF *see* corticotrophin-releasing factor
CRH *see* corticotrophin-releasing hormone
Crick, Francis, 488M
cri-du-chat syndrome, 470F
Crigler–Najjar syndrome, 349–350
CRISPR RNA (crRNA), 548
CRISPR-Cas, 547–548
cristae, 495F
crizanlizumab, 405
Crohn, Burrill, 345
Crohn disease, 121B, 300, 343, 345, 398
crotonic acids, 270F
crown (teeth), 320–321B, 321F
CRP *see* C-reactive protein
crRNA *see* CRISPR RNA
CRT *see* conformational radiotherapy
cruciferous vegetables, 527
cryostat, 535
cryotherapy, 563B
cryptococcal meningitis, 90–91
cryptococcosis, 45, 46, 90–91
Cryptococcus neoformans, 94
Cryptococcus spp., 46F
Cryptosporidium parvum, 60, 63B
crypts of Lieberkühn, 325F, 342F
crystalline silica, 525B
crystallins, 570F, 580
CT *see* computed tomography; computerised tomography; cycle threshold
CTCs *see* circulating tumour cells
ctDNA, 536
CTLA-4 *see* cytotoxic T lymphocyte antigen 4
CTLs *see* cytotoxic T lymphocytes
cubam, 331
culture, 70
curcumin, 527F
Cushing syndrome, 4, 212, 218, 220–221, 248, 249, 270, 294
cutaneous infection, 125
see also skin infections
CVA *see* cerebrovascular accident
CVID *see* common variable immunodeficiency
CVS *see* cyclic vomiting syndrome
CXC/CC chemokines, 167
cyanosis, 4T, 172, 188C, 395B, 396, 400T
cycle threshold (Ct), 74
cyclic adenosine monophosphate (cAMP), 196, 197F
cyclic AMP, 39–40, 40F
cyclic GMP (cGMP), 39
cyclic guanosine monophosphate (cGMP), 196, 197F, 232–233B
cyclic vomiting syndrome (CVS), 324B
cyclin-dependent kinase p21, 547
cyclooxygenase activity, 436B, 436F
cyclooxygenase-2 inhibitors, 520B
cyclophilin, 184F
cyclophosphamide, 563B
cycloserine, 80
cyclosporin, 183, 184F
CYP17^{a} inhibitor, 543T
CYP51 *see* cytochrome P-450 14α-demethylase

cyproheptadine, 324B
cysteine (C), 37–38, 103F, 439, 440F, 461
cystic fibrosis (CF), 328M, 454, 483, 484–486
  and cholera, 485B
  clinical features, 484T
  diagnosis and treatment, 484–486
cystic fibrosis transmembrane conductance regulator (CFTR), 483–484, 484F, 486F, 499–500
cystinuria, 461
cystitis, 60, 61F
cysts, 47F
cytarabine, 561F
cytidine, 64B, 64F
cytochrome, 293F
cytochrome *b* reductase, 332–333, 409
cytochrome b_{558}, 138
cytochrome *c*, 519B
cytochrome P-450 14α-demethylase, 91F
cytochrome P-450 aromatase (CYP19A1), 542, 551F
cytochrome P-450 systems, 362F, 363B, 363F, 365, 369–370, 370F, 373, 408M
cytochromes, 385B, 387F
cytogenetics, 466
cytokine release syndrome (cytokine storms), 103M, 185
cytokines, 100, 101–103, 107B, 123, 125, 251, 525B
  blood disorders, 384
  cachexia, 531–532B
  clinical features, 154B
  G protein-coupled receptors, 196
  general diagnosis of cancer, 531
  as hormones, 102M
  immunosuppressants/immunosuppression, 183
  immunotherapy, 543
  mast cells, 109
  nomenclature, classification and selected activities, 101–102
  superantigens and toxic shock syndrome, 124B
  see also helper T lymphocytes cells; interleukins; tumour necrosis factor
cytokinesis, 87–88, 455
cytolysins, 37–38
cytomegalovirus (CMV), 54T, 75T, 160T, 173F, 173T, 176B, 187, 414–415
cytoplasm
  apoptosis and cancer, 519–520B
  nucleocytoplasmic transport, 493–495
cytosine arabinoside, 561
cytosine (C), 71–74, 406–407, 454F, 455F, 467, 468B
  see also deoxyribonucleic acid
cytoskeletal disorders, 502–508
cytoskeleton, 456
  anthelmintic drugs, 93
  endoplasmic reticulum, disorders of, 486–487
  Golgi apparatus, disorders of the, 487–488
  Wiskott–Aldrich syndrome, 136–137
cytosol, 454–455, 519B
cytosolic phosphatase calcineurin (CaN), 183–184, 183F
cytotoxic cell-mediated immunity, 124
cytotoxic drugs, hyperuricaemia, 257
cytotoxic precursor (T_C) cells, 117, 118B, 122, 123F, 176–177
  activation, 122, 123F
  autoimmune disorders, 141
  graft rejection, 175
cytotoxic T lymphocyte antigen 4 (CTLA-4), 544
cytotoxic T lymphocytes (CTLs), 117, 122
cytotoxicity
  metronidazole, 82
  see also toxicity
cytotoxins, 60

D

dabigatran, 394B
dabrafenib, 546
dactinomycin, 540
DAF *see* decay-accelerating factor
DAG *see* diacylglycerol
dalfopristin, 75M, 78
dapsone, 395
Darling, Samuel, 46
dasatinib, 562
DAT *see* direct antiglobulin testing
dATP, 182
daunorubicin, 540, 561F
DCE genes, 165–166
DCIS *see* ductal *in situ* carcinoma
DCLRE1C (Artemis) gene mutations, 132
De Curtorum Chirugia per Insitionem (*On the Surgery of Mutilation by Grafting*) by Tagliacozzi, 174B
deacetylation, 547
death receptors (DRs), 520B
death-inducing signal complex (DISC), 520B
debridement, 94
decamers, 2F
decay-accelerating factor (DAF), 140
deep mycoses, 125
deep vein thrombosis (DVT), 229B, 445, 447
defensive virulence factors, 35–36
deferasirox, 404–405, 408
deferiprone, 408
deferrioxamine mesylate, 172F, 375
defibrillators, 431–432
deflazacort therapy, 506
degenerative diseases, 5
dehydration, 244T, 252, 308, 309F
  aspirin toxicity, 367
7-dehydrocholesterol, 249, 289F
dehydroemetine, 86
dehyroepiandsrosterone sulphate (DHEAS), 226–227
deithylcarbamazine, 93F
delafloxacin, 76
delavirdine, 64B, 64F
delayed-type hypersensitivities (DTH), 153–154
deletion mutations, 9F, 401–402, 401B, 469, 470, 486M, 521, 529, 536
  22q11 deletion syndrome, 136M
  DiGeorge anomaly, 136
δ globins, 407
dementia *see* Alzheimer disease
dendritic cells, 102–103, 123, 123M, 544
dental calculus, 337
dental hygiene, 337M
deoxyguanosine, 54B, 54F
deoxyribonucleic acid (DNA), 2F, 9F, 453–455, 514M
  ageing, abnormal proteins and the error-catastrophe theory, 574
  ageing, cancer, 579–580
  ageing, gene mutations and, 575
  ageing, programmed ageing, 574–575, 575F
  antifungal drugs, 90F
  antiprotozoal drugs, 89
  Baltimore classes, 28T
  blood groups/genotyping, 168
  chromosomal mutations, 469–477
  cystic fibrosis, 484–485, 486M
  detection of hereditary diseases, 477–479
  digestion, 319
  drug receptors, 359
  embryonic/fetal haemoglobin, 388F
  fragile sites, 467
  gene factor VIII, 413M
  genomes, 459, 459M
  haemoglobinopathies, 401B
  haemophilia A, 413, 413F
  HIV and AIDS, 62B
  inhibitors of replication/transcription, 76–77
  macrocytic anaemia, 399
  methylation, 467, 468B
  metronidazole, 82
  mitochondrial DNA and its replication, 496
  molecular biology HLA typing, 180F
  molecular model, 454F
  monoclonal antibodies, 120–121B
  nitrofurantoin, 82

nuclei and nucleocytoplasmic transport, 492, 495
P-450 systems, 363B
PAGE, 485
PCR, 71–74
repair molecules, targeted therapy, 546
replication and the cell cycle, 455–458
schematic illustration, 454F
sickle cell anaemia, 404F
thalassaemias, 407–408
thrombolysis, 349M, 439M
vaccines, 67–69
see also chromosomes; methylation; mutations; next-generation sequencing; polymerase chain reaction
depolymerisation, 537, 539–540
dermatitis, 154, 395
see also skin
Descartes, René, 194B
desferrioxamine, 375, 404–405, 408
desmin, 503
desmopressin, 204, 412F
detoxification mechanisms, 365F, 369–370, 370F, 579–580
ethanol toxicity, 369–370, 370F
liver and kidneys, 361–364
paraquat, 374–375, 375F
dexamethasone, 182, 221F, 532B
DGA *see* DiGeorge anomaly
dGTP, 182
DHA *see* docosahexaenoic acid
DHEAS *see* dehyroepiandsrosterone sulphate
DHFR *see* dihydrofolate reductase
DHMA *see* dihydroxymandelate
diabetes
ageing, 570, 583C
and cataracts, 580
glycated proteins, 573
cardiac failure, 433
endoplasmic reticulum, disorders of, 487
myocarditis/viral cardiomyopathy, 436
see also types 1/2 diabetes
diabetes insipidus, 195B, 203–204, 248
diabetes mellitus, 2, 4F, 11F, 12, 60, 143, 206
acute complications, 212–213
chronic complications, 213
chronic kidney disease, 244
diagnosis and treatment, 214–215
disorders of the exocrine pancreas, 348
EFAs, 283
hypernatraemia, 248
hypomagnesaemia, 255
infections of the skin, 55
remission, 214B
see also obesity
diabetic cataracts, 197F, 213
diabetic ketoacidosis (DKA), 212–213, 212F, 213, 269, 270B, 499
hyperphosphataemia, 254
hypophosphataemia, 254
diabetic nephropathy, 213
diabetic neuropathy, 213
diacylglycerol (DAG), 196, 197F
diagnosis and treatment
acid–base balance disorders, 274
cancer, 531–548, 553, 557–558
cancer staging, 536–537, 549–552, 553, 554F, 554T, 555, 556T
cystic fibrosis, 484–486
diabetes mellitus, 214–215
Down syndrome, 478B
G6PD, 409–411
genetic diseases, 477–479
hereditary elliptocytosis, 506–508
hereditary spherocytosis, 507–508
hypoglycaemia, 216
of infectious diseases, 75–94
antibacterial drugs, 76
antiviral drugs, 75–76
inhibitors of bacterial cell wall biosynthesis, 79–81
inhibitors of DNA replication/transcription, 76–77
miscellaneous antibiotics, 82–85
surgery, 94
ischaemic heart disease, 434–435
lysosomal storage diseases, 491
mitochondrial disorders, 498–499
muscular dystrophies, 454, 478, 488, 498–499, 504–506
myocardial infarctions, 442–444
sickle cell anaemia, 404–405
thalassaemias, 407–408
thrombocytopenia, 415
thyroid hormone disorders, 208–209
varicose veins, 447
vWD, 412–413
see also drugs; laboratory tests; prevention, control and treatment of diseases; screening; vaccines, vaccination
dialkylnitrosamines, 524
dialysis, 243–244, 499
ethylene glycol toxicity, 372
diarrhoea, 59, 60, 125, 346B, 371
diarylheptanoids, 527F
diastole, diastolic relaxation, 422, 423F
diastolic blood pressure, 426, 428
DIC *see* disseminated intravascular coagulation
dicoumarol, 394B, 394F
dictyosomes, 487, 488F
diebetes mellitus, *see also* types 1/2 diabetes mellitus
diesel fumes, 555
diet, 214–215
calorie restriction, 581–582
and cancer, 526–527
cachexia, 532B
cholesterol/lipoproteins, 439–440
fiber-rich diets, 214–215, 348, 526–527
haemorrhoids, 348
hypertension, 449
microcytic anaemias, 398
prostate cancer, 557
Refsum disease, 500–501
retinoblastoma, 563B
see also obesity
diet and disease, 279–313
diet and nutrition, 279–293
dietary malabsorption, 341, 343, 344F
dietary nonhaem iron, 333
dietary reference values (DRVs), 280F
diethylcarbamazine, 75M, 93F
diethylstilbestrol, 524T
DiGeorge anomaly (DGA), 130, 136
digestion, 319–328
absorption of its products, 328–335
mouth, 320–322
small intestine, 323, 325–328
digital rectal examination (DRE), 557F, 558
digoxin, 431F, 433
dihydrofolate reductase (DHFR), 120B, 537, 538F
dihydrotestosterone, 198F
dihydroxyacetone phosphate, 370M
1α,25-dihydroxycholecalciferol (1α,25DHCC), 192, 239–240, 289
dihydroxymandelate (DHMA), 223F
dilated congestive cardiomyopathy, 435, 436
dilute urine (polyuria), 203, 245, 248
dimethylsulphoxide (DMSO), 187F
2,4-dinitriphenol (DNP), 358B, 367
dioxygen (O_2), 362F, 374–375, 385
see also haemoglobin
dipeptides, haemoglobinopathies, 402F
diphtheria, 39–40, 50C, 69M
diplococcus bacteria, 62, 65, 65F, 73–74, 79
diploid (2N) chromosomes, 453
direct antiglobulin testing (DAT), 171
disaccharides, 325, 325M, 326T
DISC *see* death-inducing signal complex
discrepancies, 168
disease-modifying antirheumatic drugs (DMARDs), 146
dissection, 445
aortic, 445–446
disseminated intravascular coagulation (DIC), 165
distal tubules, 240, 241F
diuretics, 449

divalent metal transporter 1 (DMT1), 333
DMARDs *see* disease-modifying antirheumatic drugs
DMD *see* Duchenne muscular dystrophy
DMSA *see* agent dimercaptosuccinic acid
DMT1 *see* divalent metal transporter 1
DNA *see* deoxyribonucleic acid
DNA polymerases, 54B, 496
DNA sequencing methods, 485
  see also next-generation sequencing; polymerase chain reaction
DNA topoisomerases, 86
DNP *see* 2,4-dinitriphenol
dobutamine, 430F
docetaxel (Taxotere), 540, 541F
docosahexaenoic acid (DHA), 282, 527
Doll, William, 525–526B, 528
domain of IgG, 113M
dominant alleles, 459–461, 461T
dopamine, 201–203
dopamine-secreting nerve cells, 580
double helix, 454F, 541F
  see also deoxyribonucleic acid; helices
double-stranded breaks (DSBs), 521, 529
Down syndrome, 454, 466–467, 474–475, 474F, 475F, 477–478, 479C, 561
  prenatal diagnosis, 478B
doxorubicin, 540, 563B
doxycycline, 55, 78
DRE *see* digital rectal examination
DRESS *see* drug reaction/rash with eosinophilia and systemic symptoms syndrome
DRs *see* death receptors
drug reaction/rash with eosinophilia and systemic symptoms (DRESS) syndrome, 82
drug-receptor interactions, 359–360
drugs
  ACE, 433–434, 434M, 449
  acute kidney injury, 243
  angina, 439
  anthelmintics, 92–94, 359
  anticoagulant therapy, 435–436B
  anticoagulants, 252, 283, 301M, 394B, 394F
  anticonvulsant drugs, 299–300
  antimalarial drugs, 41–42, 42F, 395, 409
  aortic dissection, 445–446
  apoptosis and cancer, 520B, 520F
  atrial fibrillation and flutter, 431–432
  blood screening, 172–173
  breast cancer treatments, 549–550
  cachexia, 532B
  chromosomes and inherited cancers, 522
  corticosteroids, 146, 180–182, 181T, 182F, 188C, 415, 437B, 506, 532B
  Crohn disease, 345
  cystic fibrosis, 486, 486F
  diabetes mellitus, 214–215
  disorders of vitamin nutrition, 299–300
  exercise stress tests, 430
  familial hypercholesterolaemia, 441–442
  G6PD, 409–410, 410T
  hepatitis, 350–352
  hypertension, 449F
  hyperuricaemia, 257–258
  hypomagnesaemia, 255
  immune checkpoint inhibitors, 544
  immunosuppressants/immunosuppression, 121B, 129, 130T, 147, 161, 180–185, 181T
  inhibitors of DNA replication/transcription, 76–77
  lysosomal storgae disorders, 491F
  metabolic alkalosis, 270, 271F
  mitochondrial disorders, 499F
  morning-after pill, 229B
  myocardial infarction, 444
  Na^+ homeostasis disorders, 247F
  NSAIDs, 146, 147
  obesity, general management, 308–309, 309F
  pituitary function disorders, 203–204
  progeria, 582B, 582F
  purine/pyrimidine analogues, 538
  repurposing, cancer, 538M
  resistance, 60, 61–62B, 89, 94
    anthelmintic drugs, 94
    in cancer, 537M
    chemotherapy, 537
    combination therapy, 75M
    infections of the urogenital system, 60
    inhibitors of bacterial cell wall biosynthesis, 80
    MDR, 60, 92
  retinoblastoma, 563B
  SAIDs, 146
  targeted therapy, 545–547
  to treat HIV/AIDS, 63–64B, 64T
  ulcerative colitis, 345
  see also antibacterial drugs; antibiotics; anticancer drugs; antiprotozoal drugs; antiviral drugs; chemotherapy; immunosuppressants/immunosuppression; toxicology
DRVs *see* dietary reference values
dry form beriberi, 298
DSBs *see* double-stranded breaks
dsDNA genome, 28T
DTH *see* delayed-type hypersensitivies
Duchenne, Guillaume-Benjamin-Amand, 505
Duchenne muscular dystrophy (DMD), 461T, 504, 504T, 505–506
duct cells, 326–327B, 327F
ductal *in situ* carcinoma (DCIS), 549, 550T, 551
ductless glands, 191F
Duffy blood group system, 166, 167
Dulbecco, Renato, 28M
duodenal ulcers, 60, 340
duodenum, 323, 325, 328
  copper absorption, 332–333
  iron metabolism, 385B
  secretin, 336
duplication mutations, 469, 470
durvalumab, 544
dust mites (*Dermatophagoides pteronyssinus*), 151F
DVT *see* deep vein thrombosis
dwarfism, 205F
dynamic mutations, 468B
dyneins, 503
dysarthria, 300, 500
dysentery, 40F, 85B, 86T
dysmorphic features, 469–470
dyspepsia, 275C
dysphagia, 337, 338, 500
dyspnoea (shortness of breath), 149, 172
dystrophin, 503F, 504

E

E2F transcription factor, 518, 520
E3 ubiquitin ligase, 518, 518M
EAR *see* estimated average requirement
ear infections, 73
early disseminated Lyme disease, 66–67
early localised Lyme disease, 66–67
early-infant pneumonia, 55
ears
  infections of, 55–56
  ototoxicity and Ménière disease, 77B
eating disorders, 358B
  see also anorexia nervosa; bulimia nervosa; obesity
Ebola haemorrhagic fever (EHF), 33
Ebola virus, 33, 69
EBUS-TBNA *see* endobronchial ultrasound-guided transbronchial needle aspiration
EBV *see* Epstein–Barr virus
ECF *see* extracellular fluid
ECGs *see* electrocardiograms
echinocandins, 89–92
Echinococcus granulosus, 55, 94
echocardiography, 429F, 433, 434, 448
ECM *see* extracellular matrix
eczema, 137, 148M, 334B, 464, 549
ED *see* erectile dysfunction

EDTA *see* ethylenediaminetetraacetate
Edwards syndrome, 474F
EFAs *see* essential fatty acids
effector caspases, 519–520B
eflornithine, 86, 86T, 88, 88F
EGF *see* epidermal growth factor
EGFR see epidermal growth factor receptor
EHEC *see* enterohaemorrhagic
EHF *see* Ebola haemorrhagic fever
EIA *see* enzyme immunoassays
eicosanoids, 125, 148–149, 152B, 196, 239–240, 251, 281, 340–341
eicosapentaenoic acid (EPA), 282, 282F, 527, 532B
elastase, 272–273B, 326–327, 348
electrocardiograms (ECGs), 428–429, 431, 433, 434, 439, 442, 443–445, 448, 450C
electrocardiography, 435
electrolyte disorders, 239, 245–256
  Ca^{2+} homeostasis disorders, 249–253
  K^+ homeostasis disorders, 248–249
  Mg^{2+} homeostasis disorders, 255–256
  Na^+ homeostasis disorders, 246–248
  P_i homeostasis disorders, 254–255
electrolytes, 433
electromyographies (EMGs), 505
electronic (e-)cigarettes, 526B
electrophoresis, 112F, 401B, 401F, 416C, 416F
  PAGE, 485
elexacaftor, 486
eliglustat, 491F
ELISA *see* enzyme-linked immunosorbent assay
elliptocytosis, 385, 506–508
eltrombopag, 415F
EMA *see* European Medicines Agency
embolisms, 431
  thrombocytosis, 415
embryo, 228F
  see also neonates
embryogenesis, 518, 519T
embryonic haemoglobin, 388
  see also fetal haemoglobin
emetine, 86
EMGs *see* electromyographies
emphysema, 273B
EMT *see* epithelial–mesenchymal transition
emulsification, 317
enamel, 58–59, 297, 301, 302B, 320–321B, 321F
endobronchial ultrasound-guided transbronchial needle aspiration (EBUS-TBNA), 555
endocarditis, 430
endocardium, 421F, 422
endocrine system disorders, 5, 191–238, 294
  adrenal cortex, 216–224
  and autoimmune disorders, 142
  causes, 199–200
  cirrhosis, 352–353
  and GIT tract secretions, 336
  growth hormone disorders, 204–206
  hormone production, 193–195
  mechanisms of hormonal action, 195–199
  pituitary gland, 200–204
  regulation of blood glucose, 210–216
  reproductive hormones, 224–233
  structural classification of hormones, 192T
  thyroid hormone disorders, 206–210
  see also hormones
endocytosis, 488, 495
endogenous agents, 2, 2M
endogenous pyrogens, 103
endomembrane-related disorders, 486–491
  endoplasmic reticulum, 486–487
  Golgi apparatus, 487–488
endometrial cancer *see* uterine cancer
endopeptidases, 326–327
endoplasmic membrane, prepoinsulin/active insulin conversions, 211F
endoplasmic reticulum (ER), 139, 211F, 486–487, 488, 492
  ageing, 570
  lysosomal storage diseases, 488, 489T
  see also rough endoplasmic reticulum
endoscopes, 534, 534F
endoscopy, 96C
endosomes, 211, 488
endothelial cells, 103, 426
endothelial lipoprotein lipase, 441F
endothelium-dervied tissue plasminogen activator, 439M
endotoxins, 37
end-stage renal failure (ESRF), 244
energy, 280
  PEM, 293–294
  restriction and ageing, 581–582
  see also homeostasis, homeostatic diseases
energy metabolism, 191, 192F
  anthelmintic drugs, 93
enlarged atrium, 429
Entamoeba histolytica, 43, 60, 86, 86T, 89
enteral hypercaloric feeding, 532B
enteric cytopathic human orphan (ECHO), 56
enteric fever, 69
enterobacteria, 54–55
Enterobacteriaceae, 39
Enterobacteriaceae, 80
Enterococcus faecalis, 60, 78, 80, 85B
Enterococcus faecium, 78
Enterococcus vermicularis, 47–48
enterocytes, 60
  calcium absorption, 332F
  energy homeostasis, 335F
  iron absorption, 333, 334F
  lipid absorption, 330F
  nucleoside and nucleotide absorption, 329, 330F
  peptide and amino acid absorption, 328, 329F
  small intestine and homeostasis, 335F
  vitamin absorption, 331, 332F
enterohaemorrhagic (EHEC) forms of *E. coli*, 59F, 60
enteroinvasive (EIEC) forms of *E. coli*, 59F
enteropathogenic (EPEC) forms of *E. coli*, 59F, 60
enterotoxigenic (ETEC) forms of *E. coli*, 59F, 60
enterotoxins, 39F
enteroviruses, 56, 487
envelopes, 27, 29F, 492, 493F, 494, 495F
enzyme immunoassays (EIA), 116B
enzyme replacement therapy (ERT), 491
enzyme-labelled anti-immunoglobulin, 535
enzyme-linked immunosorbent assay (ELISA), 67, 116B, 116F, 140, 142, 146, 533
enzymes
  acute myeloid (myeloblastic) leukaemia, 561
  and acyclovir, 54B
  ageing, Alzheimer disease, 575
  ageing, programmed ageing, 574–575, 575F
  ageing, waste products, 574
  anthelmintic drugs, 92
  anticoagulant therapy, 436B
  antifungal drugs, 91
  antigens of the ABO system, 163–164
  anthelmintic drugs, 93
  antiprotozoal drugs, 86–88
  apoptosis and cancer, 519–520B
  arsenic poisoning, 377–378, 378F
  aspirin toxicity, 367, 368F
  basophils, 109T
  blood disorders, erythrocytes, 387F
  blood groups, fucosyltransferase, 167
  blood groups, glycosyltransferase, 167F, 167M
  breast cancer susceptibility genes 1 and 2, 521
  chemotherapy, 537
    folic acid antagonists, 537
    purine and pyrimidine analogues, 538, 539F
  chronic heart failure, 433–434, 434M
  clotting factors, 390–391
  clotting pathways, 393
  combined T and B lymphocyte deficiencies, 132F
  complement, 103–107

enzymes (*continued*)
copper absorption, 332–333
cytokines, 102–103
detoxification mechanisms, 360, 362F
dietary malabsorption, 343
digestion, 319–328
dissolution of clots, 395
drug receptors, 359F
drugs to treat HIV/AIDS, 62–64B
ethanol dehydrogenase and ethnicity, 370M
ethanol toxicity, 369–371
ethylene glycol toxicity, 372F
free radicals, 573F
G6PD, 408–411
gene therapy, 548
genetic diseases, phenylalanine hydroxylase, 5
Golgi apparatus, disorders of, 488
hormone therapy, 542, 551F
hypertension, 449
immunohistochemistry, 119B, 119F
inherited gene disorders, 461–462
inhibitors of bacterial cell wall biosynthesis, 80–81
inhibitors of DNA replication/transcription, 76–77
inhibitors of mRNA translation, 79
iron absorption, 333
kidney disorders, 239–240
lactose intolerance, 342F
lead poisoning, 376–377, 377F
lipoproteins, 441F
lysosomal storage diseases, 488, 489–490, 489F, 490, 491F
methaemoglobinaemia, 395–396
methanol toxicity, 372
miscellaneous antibiotics, 82–83, 85
mitochondrial DNA, 496
mitochondrial matrix, 495–496
molecular biology of cancer, 516B
molybdenum, 293
monoclonal antibodies, 120B
mouth, 320, 321F
muscular dystrophies, 505F
myocardial infarctions, 442
nonspecific defences, 100, 101F
nucleocytoplasmic transport, 493–494
offensive virulence factors, 38–40, 40F
P-450 systems, 91F, 362F, 363B, 363F, 365, 369–370, 370F, 373, 408M
paraquat, 374–375, 375F
phagocytic defects, 138F
selenium deficiency, 307F
sickle cell anaemia, 404F
surgery, 94
thrombolysis, 349M
tumour microenvironment, 525B
tumour suppressor genes, 518
vitamin absorption, 331
Zellweger spectrum disorder, 502
see also coenzymes; *individual enzymes...*
eosin, 534–535, 534F
eosinophils, 103, 108, 109, 123, 125, 384F, 386
EPA *see* eicosapentaenoic acid
epidemics, 32–33
epidemics and pandemics, 30–32
epidemiological investigations, 69
epidemiology, 6–7
epidermal growth factor (EGF), 525B, 546
epidermal growth factor receptor 2 (HER2), 517M, 546, 549–550, 564C
epidermal growth factor receptor (*EGFR*), 533T, 536, 546
Epidermophyton spp., 45, 89
epigenetic regulators, 547
epinephrine *see* adrenaline
epistaxis, 415
epithelial cells, 123, 288–289
immune response to parasites, 125
paraquat, 374–375
viruses and cancer, 529
see also skin
epithelial tissue, 100M, 484, 514F
carcinomas, 514F
obesity and cancer, 528
see also mucous membranes
epithelial–mesenchymal transition (EMT), 525B
epitopes, 111
epoxide hydroxylase, 362F
EPSs *see* exopolysaccharides
Epstein–Barr virus (EBV), 138, 181, 415, 529–530, 529F
ER *see* endoplasmic reticulum
ErbB2 *see* epidermal growth factor receptor 2
erectile dysfunction (ED), 232B
ergocalciferol, 289
ergosterol, 89, 90F, 91
ERK *see* extracellular signal-regulated kinase
error-catastrophe theory, 574
ERs *see* estrogen receptors
ERT *see* enzyme replacement therapy
ERV *see* expiratory reserve volume
erythema migrans, 66–67
erythroblasts (megaloblasts), 399
erythrocyte casts, 143M
erythrocyte membranes, 152
erythrocyte sedimentation rates, 70
erythrocytes, 9F, 11F
acute myeloid (myeloblastic) leukaemia, 561
blood collection, 14–15, 14F
buffering and H^+ excretion, 265
capillaries, 426
chronic myelogenous leukaemia, 562
CO_2 production/transport, 263–265
cytoskeletal disorders, 504F
glutathione peroxidase activity, 307F
hereditary elliptocytosis/spherocytosis, 506, 507–508
infective endocarditis, 430
iron absorption, 333
kidney disorders, 239–240
malaria, 41–42B
rheumatoid arthritis, 144–146
transketolase activity, 304–305, 305F, 310C
water-soluble vitamins, investigation, 305
see also blood; blood clotting; blood disorders
erythromycin, 56
erythropoietin (EPO), 398F, 406, 561
Escherichia coli, 24–25, 25F, 59F, 60
bacteriophages and phagotherapy, 84B, 84F
biochemical and molecular biological techniques, 71F
defensive virulence factors, 36
enterotoxins, 39F
inhibitors of bacterial cell wall biosynthesis, 80
inhibitors of mRNA translation, 79F
lipid A virulence factor, 37F
miscellaneous antibiotics, 82F
offensive virulence factors, 36, 36F
and UTIs, 60
esomeprazole, 359
esophageal cancer, 528
cigarettes and smoking, 525–526B
ESRF *see* end-stage renal failure
essential amino acids, 283T
essential fatty acids (EFAs), 281–283, 307–308
essential nutrients, 279, 283
esterases, 328
esters, 439–441
estimated average requirement (EAR), 280F
eteplirsen, 506F
ethanal, 371, 528
ethanoate, 370
ethanol, 369–371
see also alcohol
ethidium bromide, 179
ethylene glycol, 269, 372
see also alcohol
ethylenediaminetetraacetate (EDTA), 15, 170, 252, 377F, 411
etoposide, 540, 563B
eukaryotic cells, 483
euploidy, 473, 476–477

European Medicines Agency (EMA), 545
evusheld, 121B
Ewing sarcoma, 540
excretion, 317, 318F
  carbon dioxide, 263F
  H^+ ions, 265–267, 268F
exemestane, 551F
exercise stress tests, 430
exocrine pancreas, disorders, 348–349
exogenous agents, 2, 2M
exon 51 of DD pre-mRNA, 506F
exons, 401F, 401M, 521
exophthalmos (protrusion of the eyeballs), 143, 207–208
exopolysaccharides (EPSs), 35–36
exotoxins, 37–38, 39–40
expansion mutations, 468B
expiratory reserve volume (ERV), 150B
external beam therapy, 541–542, 558
extracellular fluid (ECF), 217–218, 244–246, 246T, 263, 267
  buffering and H^+ excretion, 265–266
  Ca^{2+} homeostasis disorders, 249F
  diarrhoea, 346B, 346F
  K^+ homeostasis disorders, 249
    hyperkalaemia, 249
    hypokalaemia, 249
  Mg^{2+} homeostasis disorders, 255
  Na^+ homeostasis disorders, 246–248
  urate metabolism, 257
extracellular matrix (ECM), 139, 291, 525B, 528
extracellular signal-regulated kinase (ERK), 546
extracellular vesicles, 525B
extractable nuclear antigens (ENAs), 146
extrahepatic cholestasis, 350
extraocular Rbs, 563B
extravascular haemolysis, 170
extrinsic apoptotic pathway, 520B
extrinsic clotting pathways, 391F
eye and ear functions, Zellweger spectrum disorder, 502
eyes, infections of, 55
EZH2 inhibition, 547

F

Fab *see* fragment antigen binding
facilitated diffusion, 328
factor B (FB), 105
factor D (FD), 105
factor H, 104T, 170
factor II (prothrombin), 392–393, 393F, 394B, 411–412, 435B
factor IIa (thrombin), 392–393, 393F, 394B, 411–412
factor IX (Christmas factor), 390T, 392, 414F
factor IXa, 435B
factor P (properdin), 105
factor V (proaccelerin), 392–393, 439M
factor VII (proconvertin), 390T, 392
factor VIIa, 392, 435B
factor VIII, 160, 390T, 392, 412–413, 413F, 413M, 417C, 439M
factor X (Stuart factor), 390T, 392
factor Xa, 391–392, 391F, 392F, 435B
factor XI (plasma thromboplastin antecedent), 390T, 392
factor XIa, 435B
factor XII (Hageman factor), 140, 140B, 390T, 392, 411–412
factor XIIa, 435B
factor XIII, 390T, 391
factor XIIIa, 389F, 391F, 393
factor XIV (Fletcher factor), 390T, 392
factors *see* clotting factors; clotting pathways
FAD *see* flavin adenine dinucleotide
FADD *see* Fas-associated domain protein
$FADH_2$, 362F
  mitochondrial matrix, 495–496
faecal fat tests, 343
faecal softeners, 348
fainting, 448
fallopian tube cancer, 547
Fallot, Etienne-Louis, 395B
false positives/negatives, 17M
familial hypercholesterolaemia, 441–442, 461, 483
familial hypocalciuric hypercalcaemia, 251M
FAP *see* adenomatous polyposis familial
farmer's lung, 153F
Fas ligands (FasLs), 520B
Fas receptors, 520B
Fas-associated domain protein (FADD), 520B
FasLs *see* Fas ligands
fasting, 210
fatal familial insomnia (FFI), 26, 462–463T
fatigue, 531, 541
fats, 280F
  see also fatty acids; lipids
fat-soluble vitamins, 287–290, 331
  absorption, 332
  clinical disorders from deficiencies, 300–301
  investigation and laboratory testing, 306–307
fatty acids, 323, 327–328, 330, 335–336, 436F, 499–500, 502, 527
  anticoagulant therapy, 436B
  cachexia, 531–532B
  dietary malabsorption, 343
  disorders of the exocrine pancreas, 348
  Lorenzo's oil, 500B
  prostate cancer, 557
  see also cholesterol; essential fatty acids; free fatty acids; sphingolipids
fatty liver disease, 295B, 370–371, 487
favism, 411B
FB *see* factor B
FBN1/MFS1 gene mutations, 480C
FC (Phe-Gly) residues, 492, 493
FD *see* factor D
FDA *see* Food and Drug Administration
FdUMP *see* fluorodeioxyuridylate
feedback mechanisms, 239F
FERMT3 mutations, 139
ferredoxin, 86, 89
ferric iron, 333
ferritin, 103M, 307, 333, 385B, 385F, 398
ferrochelatase, 376–377
ferroportin, 408
ferroportin 1, 333
ferrous iron, 333
ferroxidases, 333
fertilisation and pregnancy, 228
  see also pregnant women
fetal alcohol syndrome, 136M
fetal anaemia, 170
fetal haemoglobin (HbF), 388, 389M, 400B, 401B, 401F, 404–405, 407, 408, 416C
fetalis syndrome, 407–408
fetus, 7F, 228F
  Tay-Sachs disease, 489–490, 489T, 490F
  see also neonates; pregnant women
fever, 33, 69, 172, 432B, 531
  hay fever, 109M, 120, 148, 148M, 148T, 151
  periodic fever syndrome, 115, 135B
  rheumatic fever, 432, 432B
  schistosomiasis, 46–47, 92F
fexinidazole, 88F
FFAs *see* free fatty acids
FFI *see* fatal familial insomnia
FGF *see* fibroblast growth factor
FHIT (fragile histidine triad), 468
fiber-rich diets, 214–215, 348, 526–527
fibrillation, 432
  see also atrial fibrillation and flutter
fibrin, 389F, 390T, 391, 392, 393F, 395
  see also blood clotting
fibrinogen, 389F, 390T, 391, 393F, 393M
  see also blood clotting
fibrinolysis, 439M
fibroblast growth factor (FGF), 525B
fibroblasts, 103
fibrosis, 438
FIGLU *see* formiminoglutamic acid
filarial worms, 47, 92T
filterable agents, 529

First World War, 31, 173, 358B
FISH *see* fluorescence *in situ* hybridisation
fish and EFAs, 282F, 283
Fisher system of nomenclature, 165–166, 166T
fistulae, 269, 269M
FK506-binding protein 12 (FKBP12), 184F
FKBP12, 184
flatworms, 46–47
flavin adenine dinucleotide (FAD), 87F, 223B, 223F, 256F, 285F, 304–305, 305F, 362F, 387F, 409, 410F
flavin mononucleotide (FMN), 285F
flavin-containing monooxygenase (FMO), 362F
flaviviridae, 529T
flavodoxin, 86
flavoproteins, 285
flavour and taste, 316M
flax seeds, 527
Fleming, Alexander, 80M
Florey, Howard, 80M
flow cytometry, 161, 161–162B, 187F
FLT180a gene therapy, 414
fluconazole, 91F
flucytosine, 89
5-flucytosine, 75M, 90F
fludrocortisone, 218, 219, 220F
fluid deprivation tests, 203–204, 234C
flukes, 92T, 123
fluorescence *in situ* hybridisation (FISH), 9F, 136, 136M, 535F, 549, 562
fluoresence microscopy, 187
fluoride, 291T, 293, 301, 302B, 307
fluorodeioxyuridylate (FdUMP), 538, 539F
fluoroquinolones, 76–77
fluorosis, 302B, 307
5-fluorouracil, 538, 539F
flutter, 431–432
FMN *see* flavin mononucleotide
FMR1 (fragile X mental retardation 1), 467, 468B
folic acid (folate), 286–287
  see also vitamin B_9
Folkman, Judah, 546M
follicle-stimulating hormone (FSH), 200F, 201F, 225–227, 229B, 230–231, 230F
follicular non-Hodgkin lymphoma, 543
FOLT *see* reduced folate carrier
food *see* allergies; diet; digestion
Food and Drug Administration (FDA), 92, 92M, 121B, 173M, 176B, 536, 540, 545, 548B
  Alzheimer disease, 578–579
food poisoning, 48, 360M, 379M
formaldehyde, 372
formed elements, 383T, 384F
  see also erythrocytes; leukocytes; platelets
formiminoglutamic acid (FIGLU), 305
forward grouping/typing, 168, 168M
fosfomycin, 80F
fosmanogepix, 92M
fragile sites, 467–469
fragile X syndrome, 467, 468B
fragment antigen binding (Fab), 113F
frameshift mutations, 400M
FRC *see* functional residual capacity
free cholesterol, 439–441
free fatty acids (FFAs), 213
free radicals, 289–290, 290F, 572–573
  ROS, 525, 572, 572M
  see also antioxidants
fructose, 328
fruit and vegetables, 526–527
frusemide, 252F
fucosyltransferase (FUT3), 163, 167
fulvestrant, 517M, 542–543, 549–550
functional residual capacity (FRC), 150B
fundus reflex examination, 562B
fungal hyphae, 45F
fungi, 25, 85, 129
  antifungal drugs, 89–92
  aspergilossis, 45
  candidiasis, 45–46
  cryptococcosis, 45, 46
  DiGeorge anomaly, 136
  histoplasmosis, 45, 46
  immune response, 125
  infections of the skin, 55
  inhibitors of mRNA translation, 78–79
  methaemoglobinaemia, 395
Fusarium spp., 94
Fusidium coccineum, 78–79
fusion of gametes, 476–477
fusogen (fusing agents), 120B
FUT1 gene, 163–164
FY/CD234 glycoprotein, 167

G

G banding, 465, 466F
  see also chromosomal banding
G cells, 336
G protein-coupled receptors (GPCRs), 196, 197F, 211
G proteins, 493
G1 phase of the cell cycle, 455F
  epigenetic regulators, targeted therapy, 547
  retinoblastoma susceptibility gene, 518, 520
  tumour protein 53 gene, 518
G2 phase of the cell cycle, topoisomerase inhibitors, 455F, 540
G6PD *see* glucose-6-phosphate dehydrogenase deficiency
GA *see* Golgi aparatus
GABA *see* γ aminobutyric acid; γ aminobutyric acid
galactorrhoea, 202
Galen, 30F
gallbladder, 315, 316F, 317F, 349–353
gallbladder cancer, 528
gallstones, 350F, 506
gametes, 455, 456, 458, 473
  chromosomal mutations, 469F
  euploidy, 476F
  inheritance and mutations, 459
  see also oocytes; spermatozoon; XX chromosomes; XY chromosomes
γ aminobutyric acid (GABA), 93, 371, 373
γ carboxyglutamate, 290, 394B, 394F, 435B
γ crystallin, 570F
γ globins, 404–405, 406–407
γ globulin gene promoter, 388
γ linolenic acid (GLA), 281–282
γ polypeptides, 389F
γ rays, 529, 541
γ subunits, 101F
γ-glutamyltranspeptidase (GGT), 353B
γ-linked tripeptides, 409
GAO *see* glycolic acid oxidase
gas gangrene, 40
gastric acid, 47, 60, 251, 322–323, 324B, 336, 339–340
gastric fundus, 326B
gastric inhibitory peptide (GIP), 336
gastric intrinsic factor (GIF), 323, 323M, 331F, 340, 399
gastric juice, 322–323, 323F, 336, 338, 340
gastric reflux, 359
gastric ulcers, 340
gastrin$_{17}$ and $_{34}$, 336
gastrinomas, 341
gastritis, 337, 339–340
gastroenteritis, 76, 96, 138, 308, 324B, 342, 354C
gastrointestinal tract (GIT), 380C, 513, 514
  absorption of nutrients, 328–335
  accessory organs, 315–318
    disorders, 348–353
  administering medicines, 360
  ageing, 569
  amyloid and amyloidoses, 576B
  anthelmintic drugs, 93
  Bacteroides fragilis, 36F
  Ca^{2+} homeostasis disorders, 249, 252
  cachexia, 531B
  Chagas disease, 49B
  chemical carcinogens, 524
  chemotherapy, 537

cholesterol, 440F
cirrhosis, 352–353
definitions/structure outline, 315, 316F
detection of hereditary disease, 478
detoxification mechanisms, 361–362, 362F
diabetes mellitus, 215
digestion, 319–335
disorders, 315, 336–353
essential fatty acids, 283
ethanol toxicity, 369, 371
fat-soluble vitamins, 290
Hartnup disease, 344B
helminths, 46–48, 93, 125
hormonal control, 336
hypomagnesaemia, 255
immunosuppressants/immunosuppression, 181
infectious diseases, 25, 58–60
inherited disorders of copper metabolism, 292B
investigating nutritional disorders, 303, 305–306
kidney disorders, 239–240
lead poisoning, 376
microvilli, 325F
mucus and mucins, 322B
Na^+ homeostasis disorders, 246
obesity and cancer, 528
paraquat, 375
parental feeding, 308
PEM, general management, 308
peristalsis, 322
peritoneum, 318–319
peritonitis, 320B
preventing infectious diseases, 67
regulation of blood glucose, 210F
sepsis and systemic infections, 65–66
steatorrhoea, 371
thalassaemias, 408
thrombocytopenia, 415
tumour cells, biopsy and histoloy, 534
urate balance disorders, 256
vitamin K deficiencies, 300–301
vomiting and CVS, 324B
water in diet, 283
water-soluble vitamins, investigation, 305–306
Wernicke–Korsakoff syndrome, 298B
zinc deficiencies, 302
see also colorectal cancer; large intestine; small intestine; stomach
gastrointestinal varicose veins, 380C
gastro-oesophageal reflux disease (GORD), 338–339
Gaucher disease, 462T, 490–491
Gaucher, Philippe, 490
Gaussian distribution, 10–11B, 10F
GCSF *see* granulocyte colony-stimulating factor
GDP *see* guanosine diphosphate
GDP-fucose transporter 1, 139
GE *see* genetically engineered
Gell, Philip, 147, 148T
gemtuzumab, 561
gendicine, 547M
gene sequencing, 401B
see also next-generation sequencing
gene therapy, 414, 491, 499, 506, 536, 547–548
gene-based vaccines, 69
general swabs, 13–14, 14F
genes
amplification, 529
cancer genes, 517
definitions, 453
hereditary elliptocytosis/spherocytosis, 506, 507T
modification, bacteriophages and phagotherapy, 85B
RCDP, 501T
wild-type genes, 517
see also chromosomes; mutations
genetic code, 400M
genetic diseases, 5, 453–482
chromosomal mutations, 469–477
chromosomes and the human karyotype, 465–469
detecting, daignosing and screening, 477–479
genetics definition, 454
genetics and DNA, 454–455
genotype/phenotype, 458
inherited gene disorders, 461–465
replication and the cell cycle, 455–458
see also cancer; mutations
genetic imprinting, 470
genetic variation, 458
genetically engineered (GE) species, 175, 175–176B
genome-based ageing, 571, 574–575
genomes, 28, 459, 459M, 496–497, 536
human papilloma viruses, 530F
see also somatic mutations
genomic instability, 546
genomics, 454
genotypes, 165–166, 166T, 167T, 458
gentamicin, 77B, 77F
germ line, 459
German measles, 111F, 415
see also Rubella
Gerstmann–Sträussler–Scheinker syndrome (GSS), 26, 462–463T
GFR *see* glomerular filtration rate
GGT *see* γ-glutamyltranspeptidase
GHRH *see* growth hormone-releasing hormone
GI *see* glycaemic index
Giardia lamblia, 43–44, 43F, 60F, 89
Giardia spp., 342
giardiasis, 43–44, 85, 86T
Giemsa (G) staining, 465, 466
GIF *see* gastric intrinsic factor
gigantism, 205–206, 205F
Gilbert syndrome, 349
Gilford, Hastings, 581B
gingivitis, 337
gingivitis periodontitis, 82
GIP *see* gastric inhibitory peptide
GIT *see* gastrointestinal tract
GLA *see* γ linolenic acid (GLA)
glandular fever, 415, 529
see also Epstein–Barr virus
Gleason scores, 557–558, 558T, 559T
glioblastomas, 547
globin molecules, 317
see also haemoglobin
globulins, 112F
glomerular filtration rate (GFR), 212F, 213, 242, 244, 346B, 346F
hypercalcaemia, 252
hyperkalaemia, 249
hyperuricaemia, 257
measuring, 242–243
glomerulonephritis, 243, 244
glomerulosa of the adrenal cortex, 222
glomerulus, 240, 241F, 369
ultrafiltration and reabsorption, 241–242
glossitis, 371
GLP *see* good laboratory practice
glucagon, 196, 211F, 318
glutathione-glutathione reductase systems, 86
glucocorticoid response elements (GREs), 198–199, 199F
glucocorticoids, 181T, 182F, 506, 532B
Addison disease, 218–219, 220F
gluconate-6-phosphate dehydrogenase, 87–88, 87F
gluconeogenesis, 212–213, 370, 370M
glucose, 281M, 335–336
aspirin toxicity, 368F
impaired glucose tolerance, 573
and pancreatic disorders, 318
see also blood glucose levels
glucose-dependent insulinotrophic polypeptide (gastric inhibitory polypeptide), 336
glucose-6-phosphate dehydrogenase deficiency (G6PD), 408–411, 462T
glucosyl, 491
glucosylceramide, 490

glucuronyltransferase, 349
GLUT 4 (insulin-responsive glucose permease), 211, 212F
glutamate, 394B, 394F
glutamate gated Cl^- channels, 93
glutamate oxaloacetate transaminase (GOT), 442
glutaminase, 268F
glutamine, 267, 268F, 335–336, 335F
glutathione, 294, 364–365, 365F, 375, 409, 409F, 410F, 572–573
  paracetamol toxicity, 365, 367
glutathione peroxidases, 284M, 290F, 293, 307F
glutathione reductases, 86, 285, 304–305, 305F
glycaemic index (GI), 281–282, 281M, 526
glycans, 152, 162, 163F, 192, 413F, 441F
glycated haemoglobin (HbA_{1c}), 11, 11F, 214B, 215, 573M
glycated proteins, 571, 573
glycation, 573M
glycerol, 213
glyceryltrinitrate, 434F, 435, 438–439
glycocholate, 317F
glycogen, 335–336, 367, 368F
glycogen phosphorylase, 368F
glycogenolysis, 212–213, 360M, 367
glycolic acid oxidase (GAO), 372F
glycolipids, 162, 163F
glycolysis, 89, 367
glycoproteins, 27, 36, 62B, 100F, 162, 163F, 192, 439, 533
  Duffy blood group system, 167
  Kell blood group system, 166
  Kidd blood group system, 167
glycosomes, 502B
glycosuria, 213
glycosylation, 573M
glycosyltransferase, 167F, 167M
GM_2 ganglioside *see* Tay–Sachs disease
GM-CSF *see* granulocyte-macrophage colony-stimulating factor
goitre, 107–108, 142, 155C, 207F, 209–210, 301, 302F
Golgi apparatus (GA), 139, 332–333, 486, 487
  ageing, 570
  disorders of the, 487–488
  lysosomal storage diseases, 488, 489T
Golgi reassembly and stacking proteins (GRASPs), 487–488
Golgi vesicles, 122
golgins, 488
gonad dysgenesis, 230
  see also Turner syndrome
gonadotrophin-releasing hormone (GnRH), 201F, 204–206, 225–226, 226F, 229B, 543, 558–559
gonadotrophins, 230–231
gonorrhoea, 62, 65
good laboratory practice (GLP), 18
GORD *see* gastro-oesophageal reflux disease
goserelin, 558–559
GOT *see* glutamate oxaloacetate transaminase
gout, 257–258, 257F, 302, 539–540
GPCR *see* G protein-coupled receptors
Graafian follicles, 227–228
graft versus host disease (GVHD), 132, 159, 161, 172, 175, 179, 185–187, 186F
grafts, 174–175, 444
  isografts, 175, 180–181
  rejection, 175
Gram-negative bacteria, 37, 62, 65, 65F, 73–74, 76–80, 82, 95C
Gram-positive bacteria, 37–38, 41, 62, 65, 76–79, 82
  combination therapy, 75M
  nonspecific immune defences, 100
granulocyte colony-stimulating factor (GCSF), 132–133B, 133F, 137, 154, 186
granulocyte-macrophage colony-stimulating factor (GM-CSF), 123, 124T, 544
granulocytes, 472B
granulomatous disease, 251
  see also tuberculosis
granzymes, 122, 123F
GRASPs *see* Golgi reassembly and stacking proteins
Graves' disease, 141, 142–143, 207–208, 209–210
green tea, 527
Greenland Inuits, 283
GREs *see* glucocorticoid response elements
ground rhizomes, 527F
growth factors, 196
growth hormone (GH), 5, 192, 197, 200F, 201F, 204F, 212
  disorders, 204–206
    treatment, 206
  see also human growth hormone
growth hormone-secreting tumours, 205, 206
growth hormone-releasing hormone (GHRH), 201F, 204–205
GSH *see* reduced glutathione
GSS *see* Gerstmann–Sträussler–Scheinker syndrome
GSSG *see* oxidised glutathione
GTPase, 493–494
guanine (G), 71–73, 256F, 363B, 454F, 455F, 467, 468B, 525
  see also deoxyribonucleic acid
guanosine diphosphate (GDP), 196M
guanosine triphosphate (GTP), 89, 182, 196M, 493–494
guardian of the genome (Tp53), 518
gum infections, 337
gummas, 65
gut syndrome *see* short bowel syndrome
GVHD *see* graft versus host disease
gynaecomastia, 202, 230, 352–353
gyrase-DNA, 76–77

H

H^+, 89, 263, 264–265, 264F, 326B, 328, 331, 495–496
  excretion, 265–267, 268F
  proton beam therapy, 542
  see also acid–base balance disorders
H^+-coupled folate transporter (PCFT), 331
$H^+Ketone^-$ buffers, 213
H1N1 virus, 32
H3K27 *see* histone H3 lysine 27
H5N1 influenza subtype, 31–32
HAART *see* highly active antiretroviral therapy
haem coenzymes, 551F
haem, haem molecules, 317, 318F, 333, 385, 387–388
  detoxification mechanisms, 362F
  iron metabolism, 385B
  lead poisoning, 376–377, 377F
haem oxygenase activity, 333
haemagglutination, 168
haematemesis, 339
haematocrit, 397B, 397T
haematological examination, water-soluble vitamins, 305–306
haematological indices, 396–397, 396B
haematology, 8
haematomas, 417C
haematopoietic factors, 123
haematopoietic stem cell (HSC) transplants, 405, 408
haematopoietic stem cell transplantation (HSCT), 132, 132–133B, 136–139, 172, 174T, 179, 185–188, 188C, 537, 547, 561–562
haematopoietic stem cells (HSCs), 159
haematopoietic system, 395–396
haematoxylin, 171, 534–535, 534F
haemochromatosis, 385B, 385F
haemocytometers, 161, 386F
haemodialysis, 369F
haemoglobin C (HbC), 403
haemoglobin H disease, 401B
haemoglobin H (HbH), 406, 408
haemoglobin (Hb), 4, 385, 387–388, 388F
  anaemias, 396–399
  buffering and H^+ excretion, 265
  carbon monoxide poisoning, 374F
  CO_2 production/transport, 264–265, 264F

duplication mutations, 470
embryonic/fetal, 388
genotype/phenotype, 458
glycated haemoglobin, 573M
HbA, 387–388, 388F, 400B, 401B, 401F, 405B, 416C
HbF, 388, 389M, 400B, 401B, 401F, 404–405, 407, 408, 416C
iron, 293
methaemoglobinaemia, 395–396
M-Saskatoon and hereditary MetHb, 395M
oxygen binding curve, 388F
properties, 387–388
see also anaemias; respiration; sickle cell anaemia
haemoglobinopathies, 4, 396, 400–408
nomenclature, 400B
sickle cell anaemia, 385B, 402–405
thalassaemias, 385B, 402, 405B, 406–408
haemolysins, 37–38
haemolysis, 14–15, 165, 169–170, 169F, 169T, 170, 349, 401–402
haemolytic anaemias, 300, 410–411, 506–508
haemolytic diseases of the newborn (HDN), 152, 166–168, 170, 170M
ABO system, 170M
direct antiglobulin/Kleihauer testing, 171
prophylaxis, 171
haemolytic transfusion reaction (HTR), 171
haemolysed erythrocytes, 416F
haemoperfusion, 375
haemophiliacs, 160, 172
haemophilias, 411, 413–414, 414M, 454, 460–461, 461T
Haemophilus influenzae, 56, 56M, 134
infections of the CNS, 56
infections of the ears, 56
Haemophilus spp., 404
haemopoietic stem cell transplants, 500
haemoptysis, 143–144, 555
haemorrhage, 348, 349M, 398, 413
haemorrhoids, 345, 347–348
haemosiderin, 385B, 385F
haemosiderosis, 302
haemostasis, 389–395
see also blood clotting
haemostastic disease, 436B
hairy cell leukaemia (HCL), 154B, 543, 559, 560F, 560M
halitosis, 337
HAMA *see* human antimouse antibody
hand washing, 67F
Hansen disease *see* leprosy
haplotypes, 165, 166T
haptacorrin, 323F
haptens, 111M
hapten–protein conjugates, 154
hardening of the arteries, 438
Hartnup disease, 344B
Hashimoto thyroiditis, 208
hay fever, 109M, 120, 148, 148M, 148T, 151
Hayflick limit, 516B, 574
Hb *see* haemoglobin
Hb Constant Spring, 406–407
Hb Leiden, 401–402
Hb Philly, 401–402
HBA1/*HBA2* genes, 401
HBB DNA, 404F, 408
HbF *see* fetal haemoglobin
HbH *see* haemoglobin H
hCG *see* human chorionic gonadotropin
HCL *see* hairy cell leukaemia
HDACs *see* histone deacetylases
HDN *see* haemolytic diseases of the newborn
HE *see* hereditary elliptocytosis
head and neck cancer, 121B
health, 1–2
heart, 383, 421F
angina, 435, 436B, 438–439, 449C
arrhythmias, 273B, 429, 430, 431–432, 435, 505
artificial heart valves, 430, 433F
auscultation, 422M
block, 432
blood pressure, clinical problems, 447–449
bypass operations, 433, 439, 444F
carbon monoxide poisoning, 374
cardiac failure, 432–434
control of heartbeat, 422–423, 423F
coronary thrombosis, 411, 439, 525B
cyanosis, 395B
ECGs, 428–429, 431, 433, 434, 439, 442, 443–445, 448, 450C
echocardiography, 429F, 433, 434, 448
electrocardiography, 435
murmurs, 422M
paraquat, 374
tachycardia, 143, 396, 430, 432
transplants, 437B
heart attack *see* myocardial infarction
heart disease, 310C
see also ischaemic heart disease
heart failure, 395B, 470, 505
heart valves, 433F
disorders, 437–438
heat shock proteins (HSPs), 198–199, 198M, 199F
Heberden nodes, 579F
height, 295B, 304, 458
Heinz bodies, 401–402, 402M, 409
Heinz, Robert, 402
helices, 454F, 483, 484F, 541F
platinum-based drugs, 541F
see also deoxyribonucleic acid
Helicobacter pylori, 36, 60, 73, 82, 251
gastritis, 339–340, 339F
ulcers, 340
helminths, 24, 46–48, 55, 85, 123, 342
anthelmintic drugs, 92–94
combination therapy, 75M
drug receptors, 359
flatworms, 46–47
immune response, 125
roundworms, 46, 47–48
type I hypersensitivities, 148
helper T lymphocytes (T_H), 117, 118B, 135B
activation of, 123, 124T
autoimmune disorders, 141
flow cytometry, 161–162B
rheumatoid arthritis, 144
type IV (delayed-type) hypersensitivities, 154
hepadnaviridae, 529T
heparin, 109T, 283, 394, 394F, 411, 435B
hepatic encephalopathy, 352–353
hepatic hyperbilirubinaemia, 349
hepatic jaundice, 349
hepatic portal vein, 318F, 335–336, 425
hepatic UDP glucuronyltransferase, 349
hepatitis, 30, 73, 134, 154B, 172–173, 173T, 343, 350–352, 352F, 371, 495
blood screening, 172–173
infections from blood transfusions, 172
hepatocytes, 492F
detoxification mechanisms, 362
malaria, 41–42B
hepatomegaly, 170, 186, 294
hepatosplenomegaly, 490
hepatoxicity, 83, 85
hepcidin, 408F
hephaestin, 333
HER2 *see* epidermal growth factor receptor 2
herarelin, 206
Herceptin, 550
hereditary disorders, 73, 459–461
angiodema, 140B
cancers, 521–523, 562B
heritable retinoblastoma, 562B
Li–Fraumeni syndrome, 522B, 522F
detection of hereditary diseases, 477–479
genetics, detection of, 477–479
LHON, 496F, 497–498, 497F, 499
M-Saskatoon and hereditary MetHb, 395M
retinoblastoma, 562B
and SIDS, 129, 130T
see also inheritance
hereditary elliptocytosis (HE), 506–508
diagnosis and treatment, 507–508
hereditary nonpolyposis colorectal cancer (HNPCC), 553

hereditary persistence of fetal haemoglobin (HPFH), 389M
hereditary spherocytosis (HS), 506–508
  diagnosis and treatment, 507–508
d'Hérelle, Félix, 84–85B
herpes simplex virus (HSV), 129, 337, 495, 529–530
  see also Kaposi sarcoma-associated herpes virus
heterodonts, 321B, 321F
heteroplasmy, 498, 499
heterozygous, 400, 402, 459–460, 459F, 470, 471
hexosaminidase A, 489–490, 490F
HGPRT *see* hypoxanthine guanine phosphoribosyl transferase
HHS *see* hyperosmolar hyperglycaemic syndrome
hiatus hernias, 337–339
'Hib' meningitis, 56M
HIDS *see* hyper-IgD syndrome; hyperimmunoglobulin D syndrome
high-density lipoproteins (HDLs), 310C, 440T, 441, 441M, 441T
highly active antiretroviral therapy (HAART), 64B, 75M, 76, 486, 498B
highly processed foods, 248F
high-performance liquid chromatography (HPLC), 11F, 407–408
Hill, Austin, 525B
Hiroshima bombings, 528F
hirsutism, 4, 220–222, 230–231, 235C
His-Purkinje system, 428
histamine, 106, 107F, 109T, 125, 148–149
histidine, 526
histocompatibility complex (MHC), 544
histology, 206F, 216F, 534–535
histone deacetylases (HDACs), 520B, 547
histone H3 lysine 27 (H3K27), 547
histones, 493F
histopathology, 8–9
histoplasmosis, 45, 46, 90–91
HIV *see* human immunodeficiency virus
HLA *see* human leukocyte antigen
HMMA *see* 4-hydroxy-3-methoxymandelic acid
HNPCC *see* hereditary nonpolyposis colorectal cancer
Hodgkin lymphoma, 121B, 529, 531
homeostasis, homeostatic diseases, 5, 191, 192F, 239, 289, 495–496
  electrolyte disorders, 245–256
  glucose, 210F
  and small intestine, 335–336
  see also water
homocysteine, 439, 440F
homogenous antibodies *see* monoclonal antibodies
homologous chromsomes, 473
  see also nonhomologous chromosomes
homologous recombination (HR), 521
homologous recombination repair (HRR), 546
homoplasmy, 498, 499
homozygous, 400, 402, 459–460, 459F
honeybee (*Apis mellifera*), 152F, 154C
HONK *see* hyperosmolar hyperglycaemic syndrome
hookworms, 47, 92T, 96C
hordeolums, 55
hormonal therapy, prostate cancer, 558–559
hormone replacement therapy (HRT), 228, 549
hormone response elements (HREs), 198
hormone therapy, 536, 542–543
  breast cancer, 549–550, 551–552
hormones, 5, 291, 422
  autoimmune disorders, 141
  blood clotting, 389M
  Ca^{2+} homeostasis disorders, 249, 250
  calcium intake, 290–291
  cytokines, 102M
  fat-soluble vitamins, 287–289
  and GIT tract secretions, 336
  Graves disease, 142–143
  infections of the skin, 54
  kidney disorders, 239–240, 243
  lactose intolerance, 342
  Leber hereditary optic neuropathy, 496–497
  mechanisms of hormonal action, 195–199, 198F
  monoclonal antibodies, 121B
  myocarditis/viral cardiomyopathy, 436
  phenylketonuria, 464
  production, 193–195
  thalassaemias, 408F
  see also endocrine system disorders; sex hormones; steroids
Horvitz, H. Robert, 520B
hospitals, 10–15
  collection, transport and storing of specimens, 13–14
  intensive care units, 1F
  laboratory tests, 12–13
  proton beam therapy, 542F
  requesting a clinical test, 13
  roles of clinical laboratory testing, 11–12
  types of laboratory, 8–10
  see also diagnosis and treatment; laboratory testing
hotspots, 525
houseflies (*Musca domestica*), 23F, 48
housing, 67, 569
HPFH *see* hereditary persistence of fetal haemoglobin
HPLC *see* high-performance liquid chromatography
HPV *see* human papilloma virus
HR *see* homologous recombination
HREs *see* hormone response elements
HRR *see* homologous recombination repair
HRT *see* hormone replacement therapy
HS *see* hereditary spherocytosis
HSCT *see* haematopoietic stem cell transplantation
HSPs *see* heat shock proteins
HSV *see* herpes simplex virus
HTLV III *see* human T cell lymphotropic virus type III
HTR *see* haemolytic transfusion reaction
Huggins, Charles, 542M
HuMabMouse transgenic mouse models, 121B
human antimouse antibody (HAMA), 121B
human body louse (*Pediculus humanus*), 35F
human chorionic gonadotropin (hCG), 533T
human epidermal growth factor receptor 2 (HER2), 121B
human flea (*Pulex irritans*), 35F
Human Genome Project, 517B
human growth hormone, 195, 196F
  see also growth hormone
human herpes virus 4, 54T, 529F
human herpes virus (HHV), 29, 30T, 53, 54B, 54F, 54T, 62, 63F, 119F
human immunodeficiency virus (HIV), 2–3, 3F, 25–26, 28–29, 30F, 36B, 59, 62–64B, 172–173, 173T, 337, 498B, 529–530
  antibody deficiencies, 134
  antiviral drugs, 75
  biochemical and molecular biological techniques, 73
  blood screening, 172–173
  combination therapy, 75M
  infections of the CNS, 56
  infections from blood transfusions, 172
  methaemoglobinaemia, 395
  and SIDS, 129, 130T
human karyotype, 465–469
human leukocyte antigen (HLA), 36, 141, 161, 177–180, 185, 185M, 500
  graft rejection, 175
  matching donor and recipient, 180
  molecular biology HLA typing, 180
  typing, 179–180
  see also major histocompatibility complex
human leukocyte antigen (HLA) system, 159
human lifespan, 569
human milk, 341–342
  see also breastfeeding

human papilloma viruses (HPVs), 60, 62, 68B, 530–531, 531M
simplified genome, 530F
human pinworms, 47–48
human T cell lymphotropic virus type III (HTLV III), 62B
humanised MAb, 121B, 405, 561
humoral immune response, 9, 100, 111–124, 137
Humphrey, Hubert H., 73M
huntingtin, 494F
Huntington disease, 468–469, 494–495
Hutchinson, Jonathan, 581B
Hutchison–Gilford syndrome, 581B
hybrid myelomas, 120B
hybridomas, 120–121B, 121F
hydrochloric acid (HCl), 60, 322, 322B, 323, 324B, 326B
gastritis, 339–340
ulcers, 340–341
see also stomach acid
hydrocortisone, 182, 218, 219, 220F, 532B
hydrocortisone creams, 348
hydrogen carbonate (bicarbonate), 39–40, 249, 251, 263–269, 273B, 326B, 333, 499
aspirin toxicity, 367
blood gas analysers, 274F
ethylene glycol toxicity, 372
hormones and GIT tract secretions, 336
metabolic acidosis, 269
mixed acid–base disorders, 272
respiratory acid–base disorders, 270
ulcers, 340–341
hydrogen chloride (HCl), 416C
achlorhydia, 399
see also gastric acid
hydrogen peroxide, 409
hydrolases, 488, 489F, 493
hydrolysis, 38F, 319, 348, 349F
clotting pathways, 392–393, 393F
see also digestion; enzymes
hydroxide ions, 326
hydroxychloroquine, 146
25-hydroxycholecalciferol, 307
5-hydroxyindoleacetic acid, 344B
21-hydroxylase, 221–222, 221F
hydroxylation of barbiturates, 369
hydroxylation reactions, 362F
hydroxylations, 289
4-hydroxy-3-methoxymandelic acid (HMMA), 223B
3-hydroxy-3-methylglutaryl-coenzyme A reductase (HMG-CoA reductase), 359F, 441–442, 442F
17α-hydroxyprogesterone, 222
hydroxyproline, 344B
hydroxyurea, 404–405, 404F, 415
hyperacetylation, 547
hyperacute graft rejection, 175, 176, 177F
hyperbilirubinaemia, 170, 349–350, 353
hypercalcaemia, 203–204, 249, 251–252, 300
hypercaloric feeding, 532B
hypercapnia, 271
hypercholesterolaemia, 359, 441–442, 461
hyperglycaemia, 204–205, 212–213, 348
hypergonadotrophic hypogonadism, 230
hyper-IgD (periodic fever) syndrome, 135B
hyper-IgE (Buckley–Job) syndrome, 135B
hyper-IgM syndrome, 135B
hyperimmunoglobulin D syndrome (HIDS), 115, 135B
hyperimmunoglobulin syndromes, 135B
hyperkalaemia, 213, 217, 243, 248–249
definitions, 249
hypermagnesaemia, 255–256
hypernatraemia, 246, 248
hyperosmolar hyperglycaemic syndrome (HHS), 212–213
see also hypoglycaemia
hyperparathyroidism, 251
hyperphenylalaninaemia, 11–12
hyperphosphataemia, 213, 254–255
hyperplasia, 144, 222
see also congenital adrenal hyperplasia
hyperplastic tissue, 516
hyperprolactinaemia, 230
hyperpyrexia, 367
hypersensitivities
definitions, 147
see also immunological hypersensitivities
hypersensitivity responses, 58
hypertension, 222, 244, 294, 295T, 438, 448–449
angina, 438
cholesterol/lipoproteins, 439
varicose veins, 447
hyperthyroidism, 142–143, 207–208, 209, 209F, 342
hypertriglyceridaemia, 283
hypertrophic cardiomyopathy, 436–437
hypertrophy, 429
hyperuricaemia, 83, 85, 256, 257
hypervitaminoses, 298, 301
hypoalbuminaemia, 304, 343
hypocalcaemia, 136, 249, 250–251, 250F
hypocapnia, 448
hypochromic cells, 398F
hypochromic erythrocytes, 406F
hypogammaglobulinaemias, 133
hypoglycaemia, 204, 212, 215–216, 369, 448
diagnosis and treatment, 216
hypogonadism, 230
hypokalaemia, 39–40, 221, 248–249, 369
definitions, 249
hypolactasia, 341–342
hypomagnesaemia, 255, 307
hyponatraemia, 246
definitions, 247
oedema, 246–247
hypoparathyroidism, 250
hypophosphataemia, 251, 254
hypophosphorylated pRB, 518, 520
hypopituitarism, 204
hypotension, 172, 188C, 217, 434, 448
hypothalamus, 201F, 202, 204–205, 208, 230
hypothermia, 352–353
hypothyroidism (myxoedema), 208F, 209, 209F, 294, 342
hypotonia, 476, 508C
hypouricaemia, 256
hypovalaemia, 213
hypoventilation, 270
hypovitaminoses, 297–298
hypovolaemia, 39–40, 346
hypoxaemia, 271
hypoxanthine, aminopterin and thymidine (HAT medium), 120–121B, 121F
hypoxanthine guanine phosphoribosyl transferase (HGPRT), 120B
hyrolytic enzymes, 325

I

IARC *see* International Agency for Research on Cancer
iaschaemia, *see also* cardiac ischaemia
iatrogenic, 251
iatrogenic causes, 103M
iatrogenic diseases, 2, 5
iatrogenic infections, 60
iatrogenic thrombocytopenia, 415
ibrexafungerp, 92M
IBS *see* irritable bowel syndrome
ibuprofen, 258
IBW *see* ideal body weight
IC *see* inspiratory capacity
ICIs *see* immune checkpoint inhibitors
ICTV *see* International Committee on Taxonomy of Viruses
ICVCN *see* International Code of Virus Classification and Nomenclature
IDDM *see* insulin-dependent diabetes mellitus
ideal body weight (IBW), 294, 295B
idebenone, 499F
IDIF *see* indirect immunofluorescence
idiograms, 465
idiopathic diabetes insipidus, 203–204
idiopathic disease, 3, 5, 77B
idiopathic hypercalcaemia, 252
idiopathic hypoparathyroidism, 250

idiopathic thrombocytopenic purpura (ITP), 415
IDLs *see* intermediate-density lipoprotein
IEI Committe, 130
IEIs *see* inborn errors of immunity
IFN *see* interferon
ifosfamide, 563B
IFs *see* intermediate filaments
IgA *see* immunoglobulin A
IgE *see* immunoglobulin E
IGF-1 *see* insulin-like growth factor-1
IGF-2 *see* insulin-like growth factor-2
IgG *see* immunoglobulin G
IgM *see* immunoglobulin M
Igs *see* immunoglobulins
IIRC *see* International Intraocular Retinoblastoma Classification
ileum, 323, 325
Illumina sequencing, 517B
imatinib, 522, 523F, 562
IMDPH *see* inosine monophosphate dehydrogenase
IMDs *see* inherited metabolic disorders
immune adherence, 170
immune checkpoint inhibitors (ICIs), 544
immune system, immune response, 27, 28, 36, 36B, 99–128, 129–157
  amylodoisis/sarcoidosis, 437B
  anthelmintic drugs, 92
  autoimmune disorders, 141–147
  blood disorders, 383
  cell-mediated immune response, 9, 100, 111, 118M, 123, 124, 125, 185
    antibody deficiencies, 134
    autoimmune disorders, 141, 144
    checkpoint inhibitors, 544
    classical WAS, 137
    DiGeorge anomaly, 136
    gastritis, 340
    graft rejection, 176–177
    immunological hypersensitivities, 153
    immunosuppression, 185
    rheumatoid arthritis, 144
  complement, 103–107
  cytokines, 100, 101–103
  definitions, 99
  fungi, 125
  helminths, 125
  humoral immune response, 9, 100, 111–124, 137
  immunological hypersensitivities, 147–154
  infections of the skin, 55
  leukocytes, 103, 108–110
  nonspecific defences, 100, 101F
  PEM, 303–304
  primary immunodeficiency diseases, 130–140
  protozoa, 125
  and self–damage, 118M
  specific immune response, 110–111
  toxoplasmosis, 44
  types of immune defence, 99–100
  see also allergies; autoimmune disorders; cancer
immunoassays, 115–116B, 121B
immunodeficiency diseases, 129, 160
immunofluorescence, 118B
immunogens, 110–111
immunoglobulin A (IgA), 40, 57, 112T, 114–115, 134, 135B, 137
immunoglobulin D (IgD), 112T, 115, 135B, 135F
immunoglobulin E (IgE), 109, 112T, 115, 120, 123, 124T, 125
  allergic asthma, 149
  hyperimmunoglobulin syndromes, 135B
  mediators of inflammation, 149F
  type I hypersensitivities, 148–149, 149F
  WAS gene mutations, 137
immunoglobulin G (IgG), 95C, 104, 112–114, 112T, 117–118, 121B, 125C, 133–134, 133F, 136
  ABO blood group, 164–165
  complement and blood transfusions, 169F
  Duffy blood group system, 167
  haemagglutination, 168
  HDN, 170, 171
  Kell blood group system, 166
  Kidd blood group system, 167
  macrocytic anaemia, 399
  rheumatoid arthritis, 144–146
  thalassaemias, 408
  type I hypersensitivities, 148–149, 149F
  type III hypersensitivities, 153
  ulcers, 340
immunoglobulin M (IgM), 104, 112T, 114, 117–118, 125C, 134, 137
  ABO blood group, 164–165
  complement and blood transfusions, 169F
  haemagglutination, 168
  hyperimmunoglobulin syndromes, 135B
  Kell blood group system, 166
  Kidd blood group system, 167
  Lewis blood group system, 167
  type II hypersensitivies, 152
immunoglobulin replacement therapy (IRT), 134
immunoglobulin superfamily, 113M
immunoglobulins (Igs), 9F, 112, 131T, 160
  PEM, 303–304
immunohistochemical techniques, 142
immunohistochemistry, 118–119B
immunological disorders, 5, 181
immunological hypersensitivities, 129, 147–154
  allergic asthma, 149–151
  allergic rhinitis or hay fever, 151
  anaphylactic shock, 148T, 152B
  classification, 147, 148T
  definitions, 147
  food allergies, 151
  type I hypersensitivities, 148–149
  type II hypersensitivies, 152
  type III hypersensitivies, 152–153
  type IV (delayed-type) hypersensitivies, 153–154
immunological memory, 111
immunological rejection of transplants, 174–175
immunological tolerance, 141, 141B, 164–165
immunology, 8, 9, 99
immunoreactive trypsin (IRT), 486
immunosuppressants/immunosuppression, 121B, 147, 161, 180–185
  examples and uses, 181T
  and SIDS, 129, 130T
immunotherapy, 542, 543–545
impaired glucose tolerance, 573
inborn errors of immunity (IEIs), 129, 130T, 137, 140, 141, 454
incidence of disease, 6
incisors, 321B
incomplete penetrance, 467
incubation periods, 3
indicator media, 70, 70B
indirect immunofluorescence (IDIF), 146, 147T
indomethacin, 258F
induction therapy, 181, 186
industrial solvents, 555, 561
infantile Gaucher disease (GD2), 490
infantile Refsum disease (IRD), 500, 502
infarct definitions, 442
infarctions
  definitions, 442
  see also myocardial infarctions
infections, infectious diseases, 5, 24–25, 24M, 36–37
  anaemias, 398
  in antiquity, 24B
  course of an infection, 25
  deaths from infections, 25–26
  factors that contribute to SIDS, 129, 130T
  of the gastrointestinal tract, 58–60
  of the gums, 337
  PEM, 294
  of the respiratory system, 57–58
  of sense organs and the CNS, 55–56
  of the skin, 53–55
  transfusion and transplantation, 172
  treatment, 75–94

of the urogenital system, 60–65
viral infections, 28–30
viral replication cycle, 28, 29F
see also immune system, immune response; *individual bacterial species...*
infective endocarditis, 430
infertility, 221–222, 231–232, 297
inflammasome, 102–105
inflammation, inflammatory response, 54–55, 99F, 106, 488
acute pancreatitis, 348
alcohol and cancer, 528
anaemias, 398
angina, 439
apoptosis and cancer, 520B
cachexia, 532B
Crohn disease, 345
Duffy blood group system, 167
gonorrhoea, 62, 65
haemorrhoids, 348
obesity and cancer, 528
pathogenesis, 3
sepsis and systemic infections, 65–67
TRALI, 171–172
see also allergies; anti-inflammatories; cytokines; immune system, immune response
inflammatory bowel disease, 487
inflammatory mediators, allergies, 334, 335F
inflammatory myofibroblastic tumour, 495
influenza, 29F, 31–32, 36B, 57, 69M, 354C
Ingram, Vernon, 403B
inhalers, 151F
inherited gene disorders, 461–465
inherited metabolic disorders (IMDs), 270B
inhibitors
of bacterial cell wall biosynthesis, 79–81
of DNA replication/transcription, 76–77
of mRNA translation, 77–79
initiation (tumourigenesis), 516F, 524
injection therapy, 447
inosine monophosphate dehydrogenase (IMPDH), 182
inositol triphosphate (IP_3), 196, 197F
insertion mutations, 521
inspiratory capacity (IC), 150B
inspiratory reserve volume (IRV), 150B
insulin, 5F, 192, 193F, 197, 211F, 212, 216, 318
see also diabetes mellitus
insulin receptor (IR), 211, 483
insulin receptor superfamily, 545–546
insulin resistance, 532B
insulin-dependent diabetes mellitus (IDDM), 211
insulin-like growth factor-1 (IGF-1), 204–205, 204F, 528, 543, 557
insulin-like growth factor-2 (IGF-2), 216F
integrins, 139, 519–520B
intelligence quotients (IQ), 502
intensive care units, 1F
interferon-α (IFN-α), 102F, 138, 154B, 352, 415, 543
interferon-β (IFN-β), 102F, 154B
interferon-γ (IFN-γ), 102, 103M, 139, 154, 531B, 543
interferon-λ (IFN-λ), 102
interferons (IFNs), 102–103, 122, 123, 124T
interleukin-α, 102–103
interleukin-β, 102–103
interleukins (IL), 41, 101–103, 103M, 106, 107B, 110, 120F, 122–125, 124T, 125F
apoptosis and cancer, 520B
cachexia, 531–532B
hyperimmunoglobulin syndromes, 135B
immunosuppressants/immunosuppression, 183, 185
immunotherapy, 543
rheumatoid arthritis, 144
intermediate filaments (IFs), 502–503
intermediate-density lipoproteins (IDLs), 440F
internal radiation therapy, 541, 542
internalisation of influenza, 29F
International Agency for Research on Cancer (IARC), 524T, 526, 528
International Code of Virus Classification and Nomenclature (ICVCN), 27
International Committee on Taxonomy of Viruses (ICTV), 27–28, 27T, 28T, 529T
International Intraocular Retinoblastoma Classification (IIRC), 563B
International Society for Blood Transfusion (ISBT), 162
International Union of Immunological Societies (IUIS), 129, 130T
interphase, 455, 456F, 457F
intervention studies, 7
intestinal infections, 59–60
intolerance, 334–335B
intra/interchromosomal translocations, 471F
intracellular adhesion molecules, 29T
intracellular fluid (ICF), 244–245, 244F, 246T, 248–249, 248F, 346B, 346F
hyperkalaemia, 249
hypokalaemia, 249
intraheptic gallstone obstruction, 350
intraocular Rbs, 563B
intrathoracic lymph node assessment, 555
intravascular haemolysis, 401–402
intravenous thrombolytic therapy, 435–436B, 439M, 443–445
intrinsic apoptotic pathway, 519–520B
intrinsic clotting pathways, 391–392, 391F
intrinsic kidney injury, 243
introns, 401F, 401M, 521
invasins, 36–37
inversion mutations, 469, 470–471, 529
iodide, 291T
iodine, 301, 307, 333
iodine radiation, brachytherapy, 558
Ion Torrent sequencing, 517B
ionizing radiation, 536
ipilimumab, 544
IRD *see* infantile Refsum disease
Ireg1 membrane protein, 333
iridium treatment, 558
irinotecan, 540
iron, 291T, 293, 302, 307, 333, 334F, 343, 385, 398F, 408, 409
ferritin, 103M, 333, 385B, 385F, 398
FeS centres, 256F
metabolism, 385B
molecular model of ferritin, 385F
overload, 172
oxidation, G6PD, 409
see also haemoglobin; sickle cell anaemia
irradiation, 528–529
see also radiation; radiation therapy; ultraviolet radiation
irritable bowel syndrome (IBS), 345–346
IRT *see* immunoglobulin replacement therapy; immunoreactive trypsin
IRV *see* inspiratory reserve volume
ISBT *see* International Society for Blood Transfusion
ischaemic heart disease, 434–435, 439
diagnosis and treatment, 434–435
familial hypercholesterolaemia, 441
ischaemic stroke, 446F
islets of Langerhans, 212, 216, 318
isografts, 175, 180–181
isohaemagglutinins, 152, 164–165
isoleucine, 344B
isoniazid, 75M, 80–81, 81F, 83, 299F
isovaleric acids, 270B, 270F
ITGB2 gene mutations, 139
ITP *see* idiopathic thrombocytopenic purpura
IUIS *see* International Union of Immunological Societies
ivacaftor, 486F
ivermectin, 75M, 93
Ixodes spp., 48F, 66

J

Jacob syndrome, 475–476
Japan, Hiroshima & Nagasaki bombings, 528F
jaundice, 349, 349M, 352–353, 354C, 380C, 385, 409–410, 506

jejunoileal bypass surgery, 308–309
jejunum, 323, 325, 328, 329F
  Crohn disease, 345
  iron metabolism, 385B
  secretin, 336
Jenner, Edward, 68–69M
juvenile Gaucher disease (GD3), 490

K

kallikrein, 392
kallikreins, 390T, 392
Kallman syndrome, 230, 230M
Kaposi sarcoma (KS), 62B, 129F, 154B, 181, 529–530, 540, 543
Kaposi sarcoma-associated herpes virus (KSHV), 181, 529–530
karyograms, 465, 466F, 470F, 479C
karyotypes, 465–469, 473–474, 474F, 476F
  definitions, 465
karyotyping, 466, 477–478, 478B, 522F, 523F, 560, 562
Katsusaburo, Yamagiwa, 524
Kell blood group system, 166, 170
Kennedy disease, 468–469
keratinocytes, 103
keratins, 503
kernicterus, 170, 349
Keshan disease, 302
ketoacidosis, 211, 212, 212F, 213, 254, 499
  see also diabetic ketoacidosis
ketonaemia, 213
ketone bodies, 265F, 335–336
ketone body formation, 265
ketonuria, 213
Ketostix testing, 310C
Kidd blood group systems, 166, 167
kidney cancer
  anti-angiogenic drugs, 547
  cigarettes and smoking, 525–526B
  and obesity, 528
kidney stones, 243
kidney tubule cells, 267, 268F
  see also distal tubules; proximal tubules
kidneys, 12, 39–40, 311C
  angiotensin-converting enzyme, 434M
  anticoagulant therapy, 436B
  aspirin toxicity, 367, 369
  blood pressure, clinical problems, 448
  buffering and H^+ excretion, 265–267, 267F, 268F
  chronic renal disease, 203–204
  cryptococcosis, 46
  detoxification mechanisms, 361–364
  diabetic nephropathy, 213
  disease, 243–244
  electrolyte disorders, 245–256
  ethanol toxicity, 369
  ethylene glycol toxicity, 372
  Hartnup disease, 344B
  homeostasis, 239, 245–256, 289, 335–336
  hyperkalaemia, 249
  hypertension, 449
  immunosuppression, 180–181
  inherited gene disorders, 461
  nephrotic syndrome, 246, 246M
  paraquat, 374
  polyuria, 203, 245, 248
  renal function tests, 242–243
  SLE, 147T
  structure, 240, 241F
  thalassaemias, 406
  ultrafiltration and reabsorption, 241–242
  urobilinogen, 318F
  Zellweger spectrum disorder, 502
Kikunae, Ikeda, 316M
kinases
  anaplastic lymphoma kinase, 536, 545–546
  ataxia telangiectasia-mutated, 521, 536
  epigenetic regulators, targeted therapy, 547
  PI3K, 517M, 518T, 522, 545–546
  see also anaplastic lymphoma kinase; ataxia telangiectasia-mutated kinase
kinesins, 503
kininogen, 390T, 392
kininogenin, 392
Klebsiella pneumoniae, 60
Kleihauer testing, 171
Klinefelter syndrome, 230, 475–476
Knudson, Alfred, 521M
Köhler, George, 120B
Koichi, Ichikawa, 524
Korotkoff sounds, 427B
Korsakoff psychosis, 298B
KS *see* Kaposi sarcoma
KSHV *see* Kaposi sarcoma-associated herpes virus
Kupfer cells, 317
Kussmaul respiration, 269
kwashiorkor, 293T, 294F, 294M, 304
kynurenine, 286, 306F, 344B

L

laboratory testing
  aspirin toxicity, 368–369
  blood banks, 159F, 160
  blood groups, 168
  blood screening, 172–173
  chemical carcinogens, 526
  cirrhosis, 353, 353B
  counting leukocytes, 386B
  dietary malabsorption, 343
  disorders of the exocrine pancreas, 348–349
  evaluation, 15–17
  growth hormone disorders, 206
  lung volumes and pulmonary function tests, 150B
  lymphocyte function, 131B
  nutritional disorders, 303–307
  paracetamol toxicity, 366–367
  phagocytic defects, 138–139
  quality of clinical/pathology test results, 17–19
  rheumatoid arthritis, 145F
  SLE, 147T
  thyroid hormone disorders, 209T
  type I hypersensitivities, 149
  type IV (delayed-type) hypersensitivities, 153–154
  see also diagnosis and treatment; *individual imaging techniques...*
lactase, 342F
lactate accumulation, aspirin toxicity, 367
lactate acidosis, 499
lactate dehydrogenase (LDH), 372F, 442, 443B
lactate production, 372F
lactate secretion, 532B
lactoferin, 57
lactose intolerance, 341–342
LADs *see* leukocyte adhesion deficiencies
LAG-3 *see* lymphocyte activation gene-3
laminae, 492F, 503F
Landsteiner, Karl, 162M, 165
Langerhan cells, 123, 123M
  see also islets of Langerhans
Langerhans cell lineages, 490
lansoprazole, 359
lapatinib, 550
large cell carcinomas, 555
large granular lymphocytes (LGLs), 109, 110, 112, 113, 124
large intestine, 336
  disorders, 345–348
  junction with the small intestine, 336F
laryngeal cancer, 528
  cigarettes and smoking, 525–526B
late disseminated Lyme disease, 66, 67
latency, 29, 30T
latent periods, 3
LAV *see* lymphadenopathy virus
LDH *see* lactate dehydrogenase
LDLs *see* low-density lipoprotein
lead, 333
  poisoning, 375, 376–377
Leber hereditary optic neuropathy (LHON), 496F, 497–498, 497F, 499
lectin/MBL pathway of complement, 104–106, 105F

lectins, 36, 131B, 466
Legionnaires' disease, 487
Legionella pneumoniae, 58
Legionella pneumophila, 487
Leigh syndrome (LS), 498T, 508C
Leishmania major, 44F
leishmaniasis, 43, 44, 48, 69M, 85, 86T, 89
leprosy, 49B, 49T, 55, 76, 96, 359
leptin, 296B
  biosynthesis, 532B
Leptospira spp., 55
lesions, 4
  skip lesions, 345
leucine, 344B
leukaemias, 9F, 154B, 299, 471–472B, 514, 534, 559–563
  acute lymphoblastic leukaemia, 5, 522, 559–560
  acute myeloid leukaemia, 188C, 495, 520B, 539T, 559–561
  acute promyelocytic leukaemia, 471–472B, 561M
  chronic lymphocytic leukaemia, 9F, 559, 560
  chronic myelogenous leukaemia, 495, 522F, 523F, 543, 559, 562
  cigarettes and smoking, 525–526B
  general treatment of, 536–537
    antimicrotubular agents, 540
    topoisomerase inhibitors, 540
  hairy cell leukaemia, 154B, 543, 559, 560F, 560M
  immunotherapy, 543
  radiation, 528
  see also acute myeloid leukaemia
leukapheresis, 173M, 186
leukocoria, 562B
leukocyte adhesion deficiencies (LADs), 138–139
leukocytes, 9F, 103, 108–110, 126C, 383T, 562
  blood disorders, 383F, 383T, 384F
  blood transfusions/products, 160–161, 160T
  DAT/Kleihauer testing, 171
  haematopoietic system, 395–396
  HTR, 171
  mononuclear leukocytes, 109–110
  polymorphonculear leukocytes, 108–109, 109T, 386
  sickle cell anaemia, 405
  TME, 525B
  types, 386
  see also immune system, immune response; lymphocytes; monocytes
leukocytosis, 564C
leukodepleting/leukodepletion, 160T, 161, 171
leukotriene fatty acids, 125
leukotrienes, 125, 148–149, 152B, 281
leuprorelin, 558–559
Levey–Jennings charts, 18F
Lewis blood group system, 166–167
Lewy bodies, 580, 581M
Lewy, Friedrich, 581M
Leydig cells of the testes, 500
LFA3 *see* lymphocyte function-associated antigen-3
LFTs *see* liver function tests
LGLs *see* large granular lymphocytes
LGMDs *see* limb-girdle muscular dystrophies
LHON *see* Leber hereditary optic neuropathy
LHRH *see* luteinizing hormone-releasing hormone
lidocaine, 395
life expectancy, 569
Li–Fraumeni syndrome (LFS), 522B
ligases, 521
limb-girdle muscular dystrophies (LGMDs), 504T
lincomycin, 78
lincosamides, 78
linezolid, 61B
lingual lipase, 316M, 320, 323
lipases, 308–309, 319, 322–323, 323F, 327–328, 327F, 328M, 441F
  phospholipases, 38F, 76
  see also gastric juice; pancreatic juice
lipid A virulence factor, 37F
lipid peroxidation, 375F, 528
lipid peroxides, 290F
lipid vitamin derivatives, 192
lipids, 280F, 281–283, 310C, 317
  absorption, 330
  breakdown, cachexis, 531–532B
  flavour and taste, 316M
  see also fatty acids; phospholipids
lipoate, 331F
lipofuscins, 574F
lipolysis, 531–532B
lipopolysaccharide (LPS), 105F, 131T
lipoprotein particles, 439–440
lipoproteins, 439–442
  see also cholesterol
liposomes, 486M
liposuction, 309
lipoxygenase activity, 93
liquid biopsy, 536
liraglutide, 308–309
LISS *see* low ionic strength saline
Listeria monocytogenes, 111
Listeria spp., 34
Listeriae, 36–37
lithium, 203–204
  poisoning, 375
liver, 70, 311C, 315–318, 349–353, 380C
  anticoagulant therapy, 436B
  blood glucose regulation, 210
  cholesterol, 440–441
  clotting pathways, 394B
  deep vein thrombosis, 447
  detoxification mechanisms, 361–364
  erythrocytes, 385
  ethanol toxicity, 369–371
  homeostasis, 335–336
  iron metabolism, 385B
  LSDs, 490–491
  paraquat, 374
  prevention of clotting, 394
  prothrombin, 394B
  thalassaemias, 408
  Zellweger spectrum disorder, 502
  see also bile
liver cancer
  and alcohol, 528
  anti-angiogenic drugs, 547
  cachexia, 531–532B
  cigarettes and smoking, 525–526B
  and obesity, 528
liver disease, 246, 352
  fatty liver, 295B, 370–371, 487
  see also cirrhosis
liver function tests (LFTs), 343, 351B, 353B, 380C, 433
liver peroxisomes, 499F
liver transplants, 353, 371
LMNA gene mutations, 581–582B
lobules, 317, 361F
lonafarnib, 582B, 582F
loops of Henle, 240, 241F
Lorenzo's oil, 500, 500B
low ionic strength saline (LISS), 168
low-density lipoproteins (LDLs), 289–290, 440F, 440T, 441, 441T, 461, 483
lower reference nutrient intake (LRNI), 280F
LPS *see* lipopolysaccharide
LRNI *see* lower reference nutrient intake
LS *see* Leigh syndrome
LSDs *see* lysosomal storage diseases
lumacaftor, 486
lumen, 315, 316F, 328
lung cancer, 121B, 247, 514, 533, 534F, 555–556
  antimicrotubular agents, 540
  cachexia, 531B
  cellular image, 513F
  chemical carcinogens, 524
  cigarettes and smoking, 525–526B, 555
  fragile sites, 468
  molecular diagnosis, 536
  NSCLC, 121B, 518T, 523T, 533T, 536, 539T, 545–547, 545T, 555–556
  nucleocytoplasmic transport, 495
  platinum-based drugs, 541

lung cancer (*continued*)
risk factors, 555
symptoms, diagnosis and staging, 536, 555, 556T
TME, 525B
topoisomerase inhibitors, 540
treatment, 556
see also smoking
lungs
antiglomerular basement membrane disease, 143–144
asthma, 45, 109M, 120, 148, 148M, 148T, 149–151, 271, 273B, 275C
buffering and H^+ excretion, 265–267, 268F
carbon monoxide poisoning, 374
CO_2 production/transport, 264F
COPD, 149, 150B, 272–273B
cryptococcal meningitis, 90–91
cryptococcosis, 46
cyanosis, 395B
farmer's lung, 153F
Gaucher disease, 490
histoplasmosis, 46
hypertrophic cardiomyopathy, 436–437
immune response to parasites, 125
Leber hereditary optic neuropathy, 496–497
methaemoglobinaemia, 395
paraquat, 374–375
pulmonary arteries, 424–425
pulmonary circulations, 422
pulmonary oedema, 172, 434, 436–437
SCLC, 555
smoking, 273B, 275C, 310C, 363B, 496–497, 525–526B, 555, 561, 564C
thrombocytosis, 415
TRALI, 171–172
volumes and pulmonary function tests, 150–151B
luspatercept, 408
luteinizing hormone (LH), 200F, 201F, 225–227, 229B, 230–231, 230F
luteinizing hormone-releasing hormone (LHRH), 542–543, 543T
lycopene, 527F, 557
Lyme disease, 48, 66–67, 73
lymphadenopathy virus (LAV), 62B
lymphatic filariasis, 93
lymphocyte activation, 118F
lymphocyte activation gene-3 (LAG-3), 544
lymphocyte function, 131–132B
lymphocyte function-associated antigen-3 (LFA3), 185
lymphocytes, 386
ITGB2 gene mutation and CD18 expression, 139
LGLs, 109, 110, 112, 113, 124
serological HLA typing, 179–180
Wiskott–Aldrich syndrome, 136–137
lymphoid stem cells, 110
lymphomas, 514, 529–530, 540, 543, 552T
lymphopenia, 304
Lynch syndrome, 553
lysins, 37–38
lysogenic/temperate bacteriophages, 84
lysosomal storage diseases (LSDs), 488–491, 489T
diagnosis and treatment, 491
lysosomal trafficking regulator protein (LYST), 137
lysosomes, lysosomal enzymes, 94, 108, 488, 491, 574, 574F
lysozyme, 100, 100M, 101F, 322B, 325
LYST *see* lysosomal trafficking regulator protein
lytic/virulent bacteriophages, 84

M

M phase *see* meiosis; mitosis
MA *see* megesterol acetate
MAbs *see* monoclonal antibodies
MAC *see* membrane attack complex; microphthalmia, anophthalmia and coloboma
McCarthy, John, 548
macrocytic anaemias, 396–397, 397T, 399
macrolides, 78, 183, 184F
macrophages, 28, 58, 102–103, 106, 109–110, 122, 123, 124, 294, 385
ageing, 574F
apoptosis and cancer, 520B
graft rejection, 175
ITGB2 gene mutation and CD18 expression, 139
TME, 525B
macupuscular skin rashes, 186M
magnesium (Mg^{2+}), 40F, 82, 291T, 307, 388F, 458F
distribution in the body, 255F
homeostasis disorders, 255–256
magnetic resonance imaging (MRI), 7, 70, 95C, 434, 435, 446, 578F
AI in oncology, 548B
Alzheimer disease, 577–578, 578F
Creutzfeldt–Jakob disease, 463F
muscular dystrophies, 505–506
oncology, artificial intelligence, 548B
pineal gland, 195B
pituitary function, 202
prostate cancer, 557, 558
radiation therapy, 541
retinoblastoma, 562–563B
stroke, 446
tumour detection/localisation, 533
Wernicke–Korsakoff syndrome, 298B
maintenance therapy, 181, 186
major histocompatibility complex (MHC), 36B, 122F, 124B, 161, 177–180, 185
activation of T_C cells, 122F
activation of T_H cells, 123F
alloimmunisation, 172
graft rejection, 175, 177B
superantigens and toxic shock syndrome, 124B
TRALI, 171–172
see also human leukocyte antigen
malabsorption, 208–209, 300, 342, 348, 531B
malaria, 25–26, 36B, 41–42B, 69M, 85, 86T
Duffy blood group system, 167
G6PD, 409, 409M
sickle cell anaemia, 405, 405B
see also antimalarial drugs
malignant disease, 251–252, 456M, 514
see also cancer
malnutrition, 129, 130T, 208–209, 303T, 398
MALTs *see* mucosa-associated lymphoid tissues
mammography, 522B, 534, 549
mandrake plant (*Madragora*), 540
manganese (Mn^{2+}), 291T, 293, 302
mannans, 125
mannose-binding lectin (MBL), 140F
mannose-6-phosphate, 488
Mantoux reaction, 153
MAO *see* monoamine oxidase
MAPK *see* mitogen-activated protein kinase
marasmus, 293T, 294
Marfan syndrome (MFS), 480C
marker proteins, 118–119B, 118T, 119F
see also biomarkers
MASP-2 *see* MBL-associated serine protease
mast cells, 103, 106, 107F, 109, 123, 125, 148–149, 149F, 384F
mastectomy, 521, 551
matched sibling transplants, 136
maternal haemoglobin *see* adult/maternal Hb
matrix, 27, 495–496, 495F, 499
matrix metalloproteinases (MMPs), 525B
May–Grünwald–Giemsa, 534–535
MBL *see* mannose-binding lectin
MBL-associated serine protease (MASP-2), 105
MBTPS1 *see* membrane-bound transcription factor protease site 1
MBTPS1 (16q23.3-q24.1) gene, 488
MCH *see* mean cell Hb
MCHC *see* mean cell Hb concentrations
MCSs *see* membrane contact sites
MDR *see* multidrug resistance
MDs *see* muscular dystrophies
MDSs *see* myelodysplastic syndromes

mean cell Hb concentrations (MCHC), 396, 397B, 397T, 416C
mean cell Hb (MCH), 396, 397B, 397T, 398, 416C
mean cell volume (MCV), 396, 397, 397B, 397T, 398, 399, 416C
measles, 30–31, 68, 68B, 111F, 131
measles, mumps and rubella (MMR) vaccines, 68, 68B, 131
mebendazole, 93F, 359
mecaptopurine, 182F
mechanistic target of rapamycin (mTOR), 184
medical microbiology, 8, 9–10
medicines
  administering of, 360
  see also drugs
Medieval/Renaissance period, 174B
medroxyprogesterone acetate (MPA), 532B
mefloquine, 86
megakaryocytes, 136, 384F, 562
  see also platelets
megaloblastic anaemia, 299–300, 305, 339, 399F
megaloblastic erythropoiesis, 404
megaloblasts, 416C
megesterol acetate (MA), 532B
meiosis, 455, 455F, 456–458, 476F
  euploidy, 476F
  nondisjunction of chromosomes, 473F
meisosis, 456–458
  chromosomal mutations, 469F
  deletion mutations, 470
melanin, 136–137, 137F
melanoma, 121B, 528, 529, 543
melarsoprol, 86T, 87–88
MELAS *see* mitochondrial encephalomyopathy, lactic acidosis and stroke-like episodes
melatonin, 194B, 194F
membrane attack complex (MAC), 104–106, 105F, 106F
membrane contact sites (MCSs), 486, 487, 487T
membrane-bound transcription factor protease site 1 (MBTPS1), 488
membranes
  antiglomerular basement membrane disease, 143–144
  endomembrane-related disorders, 486–491
  humoral immunity, activation of T lymphocytes, 122
  plasma membrane-related disorders, 483–486
  see also plasma membrane-related disorders
memory cells, 118
MenACWY vaccines, 68B
menaquinones (K_2), 290
Mendel, Gregor, 459, 467, 480C, 522B
Ménière disease, 77B
Ménière, Prosper, 77B
meningioma, 542
meningitis, 2F, 56, 56M, 69M, 73, 90–91, 129
meningococcal meningitis, 129
Menkes disease, 292B, 333
menopause, 225, 227, 228, 230, 549
  see also postmenopausal women
menstrual cycle, 17, 60, 225–228, 229B, 230, 235C, 310C, 385B, 398
  fertilisation and pregnancy, 228
  menopause, 228
  see also amenorrhoea; menopause
mercury, 375
  poisoning, 379
MERRF *see* myoclonic epilepsy and ragged-red fibre
MERS *see* Middle East respiratory syndrome
mesentery, 315, 319F
mesocolon, 315, 319F
messenger RNA (mRNA), 40, 454–455
  ageing, abnormal proteins and the error-catastrophe theory, 574
  antiprotozoal drugs, 86
  cancer vaccines, 544
  haemoglobinopathies, 401B
  immunosuppression, 184
  inhibitors of translation, 77–79
  intracellular receptors, 198–199, 199F
  mitochondrial disorders, 499
  next-generation sequencing, 517B
  nuclei and nucleocytoplasmic transport, 492
  3'polyadenylation, 407M
  sickle cell anaemia, 405
  thalassaemias, 406–407
  vaccines, 69
metabolic acid–base disorders, 269–270
metabolic acidosis, 243, 269, 372, 395
metabolic alkalosis, 270
metabolic syndrome, 215B
metabolism
  acid–base balance disorders, 263F
  anticoagulant therapy, 435B
  anthelmintic drugs, 93
  buffering and H^+ excretion, 265
  copper, 292B
  ethanol, 369–370
  iron, 385B
  mitochondrial matrix, 495–496
  prostaglandins, 93
  see also biosynthesis; digestion; enzymes
metacentric chromosomes, 465F
metal ion chelator phytate, 302F
metalloproteins, 345F
metals, poisoning, 375–379
metaphase, 456F, 458, 466
metastasis, 3F, 514
  see also cancer
methaemoglobin, 387, 402
methaemoglobin reductase, 387F, 409
methaemoglobinaemia (MetHb), 395–396
methahaemoglobin, 387
methanol, 269, 372
  see also alcohol
MetHb *see* methaemoglobinaemia
methicillin, 61B
methicillin-resistant *S. aureus*, 80
methionine, 367F
methotrexate, 146, 181, 182, 183F, 537, 538F
methyl groups, 468B
methylation, 467, 468–469, 468B, 517B, 547
  alcohol and cancer, 528
1-methyl-3-carboxamido-6-pyridone, 305
5-methylcytosine (Me^5C), 468B, 468F
methyldopa, 202F
methylene tetrahydrofolate polyglutamate, 399
methylene tetrahydrofolate (THF), 538, 539F
methylmercury, 379M
1-methylnicotinamide, 305
metoprolol, 431, 449F
metronidazole, 82, 86, 86T, 89F, 340–341
metyrapone, 221F
mevalonate, 441–442, 442F
mevalonate kinase (MK), 135B, 135F
Mexican yam (*Pachyrhizus erosus*), 229B
MFS *see* Marfan syndrome; microfilaments
MG *see* myasthenia gravis
Mg^{2+} deficiencies, 250–251
MHC *see* histocompatibility complex; major histocompatibility complex
micelles, 330F
microbial virulence factors, 34–41
microbiological test strips, 70B
microcytic anaemias, 396–397, 397T, 398
microcytic erythrocytes, 406F
microcytic hypochromic anaemia, 380C
microfilaments (MFs), 502–503
microphthalmia, anophthalmia and coloboma (MAC), 488, 489T
microRNA (miRNA), 467, 467M, 525B
Microsporum spp., 45, 89
microtome, 534–535
microtubules (MTs), 93, 456, 487–488, 502–503, 537, 539–540, 541F
microvilli, 325F
Middle East respiratory syndrome (MERS), 32F
miglustat, 491F

MiHA *see* minor histocompatibility differences
milk-alkali syndrome, 251
Milstein, César, 120B
mineralocorticoids, 218–219
minerals, 290–291
  absorption, 332–333
  investigation and laboratory testing, 307
  nutritional disorders, 301–302
  recommended intakes and sources, 291, 292T
minocycline, 78
minor histocompatibility differences (MiHA), 175
miscellaneous group clotting factors, 390–391
miscellaneous/historical antibiotics, 78, 82–85
mitochondria, 453, 461, 495–499
  antiprotozoal drugs, 86
  endoplasmic reticulum, disorders of, 486–487
  free radicals, 572
  water-soluble vitamins, 285
mitochondrial disorders, 508C
  causes of, 496–497
  cristae, envelope and matrix, 495F
  diagnosis and treatment, 498–499
  DNA and its replication, 496
  examples of disease, 498T
mitochondrial DNA (mtDNA), 496–499
  acquired mitochondrial DNA defects, 498B
  and its replication, 496
mitochondrial encephalomyopathy, lactic acidosis and stroke-like episodes (MELAS), 496F
mitochondrial envelope, 495F
mitochondrial matrix, 27, 495–496, 495F
mitogen-activated protein kinase (MAPK), 546
  angiogenesis, 546
mitogens, 131, 131B
mitosis, 455, 455F, 456–458, 466, 467, 473
  apoptosis and cancer, 519–520B
  chemotherapy, 537, 540
  see also individual mitotic phases...
mixed acid–base disorders, 272
mixed function oxidases (MFOs), 362F
mixed type cancers, 514
MK *see* mevalonate kinase
MMPs *see* matrix metalloproteinases
MMR *see* measles, mumps and rubella vaccine
MNLs *see* mononuclear leukocytes
MODS *see* multiorgan dysfunction syndrome
molecular biology HLA typing, 180
molybdenum, 291T, 293, 302, 307
monoamine oxidase (MAO), 223B
monoclonal antibodies (MAbs), 120–121B, 179, 185, 188C, 533, 535, 545–547, 550, 553
  humanised, 121B, 405, 561
  see also immunoglobulins
monocytes, 93, 102–103, 109–110, 123, 384F, 386, 562
  serological HLA typing, 179–180
  TME, 525B
  TRALI, 171–172
  type IV hypersensitivities, 153–154
  see also macrophages
mononuclear leukocytes (MNLs), 108, 109–110
mononuclear phagocytic system, 317
mononucleosis, 529
monooxygenase systems, 362F
  see also cytochrome P-450s
monophospholipids, 330
monoploid fetuses, 476
monoploidy, 476
monosaccharide absorption, 328, 329F
monosodium glutamate (MSG), 316M
monosodium urate crystals, 257F
monosomy, 473
Moraxella catarrhalis, 57, 134
morbidity, 3, 138, 213
morning-after pill, 229B
morphine, 320B, 358F, 358M
mortality, 3, 213
mouse anti-CD3, 185
mouse double minute 2 (MDM2)-proteasome system, 518M
mouth, 320–322
  disorders, 337
  see also teeth
mouth cancer
  cigarettes and smoking, 525–526B
  see also oral cancer
mouth ulcers, 337, 340
MPA *see* mycophenolate
MS *see* multiple sclerosis
M-Saskatoon and hereditary MetHb, 395M
MTs *see* microtubules
mucins, 100M, 302B, 322B, 337
mucosa-associated lymphoid tissues (MALTs), 117, 120
mucosal cells, 336
  coeliac disease, 342F
  ulcers, 340
mucous membranes, 100, 100M, 288–289
mucus, 320, 322B
Mullis, Kary, 71–72
multidrug resistance (MDR), 60, 92
multifactorial disease, 2
multiorgan dysfunction syndrome (MODS), 66
multiple sclerosis (MS), 121B, 141
mumps virus, 56, 415
murine IgG, 121B
murmurs, 422M
Murray, Joseph, 173, 180–181, 181M
muscle contraction, 290–291, 503F
muscle tissue degeneration, 505
muscular dystrophies (MDs), 454, 478, 488, 504–506
  management and treatment, 505–506
mutagenic agents, detoxification, 579–580
mutant haemoglobins *see* haemoglobinopathies
mutations, 5, 132
  22q11 deletion syndrome, 136M
  adrenoleukodystrophy, 499–500
  and ageing, 575, 579–580
  blood disorders
    erythrocytes, 385
    haemoglobinopathies, 396, 400–408
  Chédiak–Higashi syndrome, 137–138
  complement deficiencies, 140
  COPD, 272–273B
  definitions, 453
  endoplasmic reticulum, disorders of, 487
  Golgi apparatus, disorders of, 488
  Hartnup disease, 344B
  hereditary angiodema, 140B
  hereditary elliptocytosis/spherocytosis, 506–508
  hyperimmunoglobulin syndromes, 135B
  inherited disorders of copper metabolism, 292B
  lysosomal storage diseases, 489, 490
  Menkes disease, 333
  methaemoglobinaemia, 395–396
  mitochondrial disorders, 496–499
  M-Saskatoon and hereditary MetHb, 395M
  muscular dystrophies, 504–505, 504T
  phagocytic defects, 138–139
  plasma membrane-related disorders, 483–486
  Refsum disease, 500
  SCID, 131
  thalassaemias, 406–407
  ultraviolet radiation, 289F, 521, 528–529, 571B
  WAS protein disoders, 136–137
  see also cancer; genetic diseases; immune system, immune disorders
myasthenia gravis (MG), 141, 144
mycobacteria, 76–77, 77F, 129, 543
  arthritis, 579
Mycobacterium abscesses, 77
Mycobacterium avium-complex, 63T, 77
Mycobacterium bovis, 80, 153
  type IV (delayed-type) hypersensitivies, 153

Mycobacterium leprae, 36–37, 55, 77
Mycobacterium tuberculosis, 25, 58F, 63T, 77F, 78F, 80, 83, 85, 111
mycophenolate (MPA), 182F
Mycoplasma genitalium, 73–74
Mycoplasma pneumoniae, 58
myelin of the CNS, 502
myeloblastic leukaemia, 495
myelodysplastic syndromes (MDSs), 516B
myeloid stem cells, 562
myeloma, 9F, 126C, 490, 543
myeloma cell-myeloma cell fusions, 120B
myeloproliferative disease, 415, 257–258
myocardial iaschaemia, 438
  see also cardiac ischaemia
myocardial infarctions, 215B, 395, 429, 432, 432M, 433, 436B, 450C
  angina, 439
  biomarkers, 442, 443B
  blood pressure, 447–448, 449
  diagnosis and treatment, 442–444
  thrombocytosis, 415
myocardial oxygen and angina, 438–439
myocarditis, 436
myocardium, 421F, 422
myoclonic epilepsy and ragged-red fibre (MERRF), 496F
myoglobin, 442–443
myosins, 503
myotonic dystrophy, 468–469
myxoedema, 208F
  see also hypothyroidism

N

N-acetyltransferase (NAT), 194B
Na^+-coupled vitamin C transporters (SVCT1 and -2), 331
Na^+/K^+-ATPase, 245–246, 245F, 326B
Na^+-dependent multivitamin transporter (SMVT), 331
Na^+-dependent transporters, 328, 329F
NAD^+, 80–81, 81F, 305
  ethanol toxicity, 370F
  structure of, 286F
NADH-cytochrome b_5 reductase, 387F, 395–396
NADH-dependent enoyl-ACP reductase, 80–81, 81F
$NADP^+$ coenzyme, 305, 408–409, 409F
  G6PD, 411B
  structure of, 286F
NADP-cytochrome b_5 reductase, 409
NADPH *see* nicotinamide adenine dinucleotide phosphate
naftifine, 89
Nagasaki bombings, 528
$NaHCO_3$, 326, 327F
Napoleon Bonaparte, 378–379B
napthalene (moth balls), 409
NARP *see* neurogenic muscle weakness, ataxia and retinitis pigmentosa
nasopharyngeal carcinoma, 529
NAT *see N*-acetyltransferase
natalizumab, 121B
National Health Service (NHS), 67F, 68B, 180F, 527, 531, 531M, 553M
  cervical screening programmes, 531M
  TNM system, 555, 556T
  xenobiotics, 358B
National Institute for Health and Care Excellence (NICE), 149, 171, 295B, 442, 557–558, 558T
National Marrow Donor Program (formerly Be the Match), 185M
National Toxicology Program (NTP), 524T
natural antibodies, 164–165
natural killer cells (NKCs), 101–103, 110, 122, 123, 124B, 131, 137, 185, 384F, 520B, 543
nature and investigation of diseases, 1–21
  characterisitc features and clinical conditions, 2–5
  classification, 5–6
  epidemiology, 6–7
  evaluating laboratory tests, 15–17
  investigating diseases, 7
  quality of clinical/pathology test results, 17–19
  summary, 20
  types of laboratory, 8–10
nausea, 532B *see* vomiting
nausea, radiation therapy, 541
NBDs *see* nucleotide binding domains
NBT *see* nitroblue tetrazolium
NCF-A *see* chemotactic factor for neutrophils
Necator americanus, 47, 51C, 92T, 96C
neck (teeth), 320–321B, 321F
necrosis, 519B
neglected tropical diseases (NTDs), 49–50B
  see also Chagas disease
Neisseria gonorrhoeae, 62, 65, 65F, 73–74, 79
Neisseria meningitidis, 2F, 36, 56, 69
Neisseria spp., complement deficiencies, 140
nematodes (roundworms), 46, 47–48, 92T, 93, 109
neonatal adrenoleukodystrophy (NALD), 499–500, 502
neonates, 300–301, 349M
  baby and chilhood vaccination programmes, 68B
  breastfeeding, 202–203, 202F, 203T
  CVS/amniocentesis, 466, 477F
  cyanosis, 395B
  cystic fibrosis, 484T
  diabetes mellitus, 216
  Down syndrome, 474F
  eye infections, 55
  graft versus host disease, 172
  HDN, 152, 166–168, 170–171, 170M
  hereditary elliptocytosis/spherocytosis, 507–508
  jaundice, 349, 349M
  Menkes disease, 292B
  number and arrangement of teeth, 320B, 320F
  PKU screening, 465, 465M
  Prader–Willi syndrome, 470
  reverse grouping/typing, 168M
  screening for hereditary diseases, 479
  stomach lipases, 323
  unconjugated hyperbilirubinaemia, 349–350
  vitamin E deficiency, 300
  vitamin K deficiency, 300–301
  see also neonates; pregnant women
neoplasia, 520B
neoplasms, 243, 514
neoplastic diseases, 5
nephelometry, 116B
nephrogenic diabetes insipidus, 203–204
nephrons, 240
nephropathy, 4, 12
nephrotic syndrome, 246, 246M
nephrotoxins, 243
neratinib, 550
nervous system (NS), 423F
  amyloid and amyloidoses, 576B
  see also central nervous system
neuroblastomas, 540
neurodegenerative disorders, 487, 488
neuroendocrine carcinoma, 555
neurogenic muscle weakness, ataxia and retinitis pigmentosa (NARP), 496F
neuroglycopenia, 215
neurotensin, 336
neurotransmitters, 291
  calcium intake, 290–291
neutrophils, 103, 106, 108, 126C, 136–137, 137F, 257, 258, 272–273B, 384F, 386
  ITGB2 gene mutation and CD18 expression, 139
  TME, 525B
  TRALI, 171–172
  type III sensitivities, 152–153
New World, 30–31
next-generation sequencing (NGS), 401B, 407–408, 478, 499, 536
  chronic myelogenous leukaemia, 562
  molecular biology of cancer, 516, 517B
  rhizomelic chondrodysplasia punctata, 502

NFATs *see* nuclear factors of activated T cells
NGS *see* next-generation sequencing
NHEJ *see* nonhomologous end-joining
niacin/nicotinic acid, 286F
  see also vitamin B_3
NICE *see* National Institute for Health and Care Excellence
nickel, 333
nicotinamide, 83F, 344B
nicotinamide adenine dinucleotide (NADH), 370, 372, 408M, 495–496, 497F
nicotinamide adenine dinucleotide phosphate (NADPH), 138, 232, 284, 286, 408, 408M
  detoxification mechanisms, 362F
  familial hypercholesterolaemia, 441–442
  G6PD, 408–411, 411B
  paraquat, 374–375, 375F
  trypanosome infections, 87–88
nicotinic cholinergic receptors, 93
NIDDM *see* noninsulin-dependent diabetes mellitus
nifurtimox, 86, 86T, 87F
night blindness, 5–6, 300
nilotinib, 522, 523F, 562
nitric oxide (NO), 232–233B, 404
nitroblue tetrazolium (NBT), 138
nitrofurantoin, 82F
nitrogen mustards, 538T
nitroimidazoles, 76
N-nitrosamines, 555
nitroureas, 538T
nitrous oxide (laughing gas), 395
nitrovasodilators, 439
nivolumab, 544
NKCs *see* natural killer cells
NO *see* nitric oxide
Nocardia spp., 138
nomograms, 366F
non-small cell lung cancer (NSCLC), 121B, 518T, 523T, 533T, 536, 539T, 545–547, 545T, 555–556
noncoding RNA, 517B
noncommunicable diseases, 24M
nondisjunction of chromosomes, 458, 473F, 475–476, 476F
nonessential amino acids, 283T
nonessential nutrients, 279
nonhaemolytic febrile transfusion reactions, 171
nonheritable retinoblastoma, 562B
nonhistone proteins, epigenetic regulators, targeted therapy, 547
non-Hodgkin lymphoma, 129
nonhomologous chromosomes, 471, 471F, 472, 473
nonhomologous end-joining (NHEJ), 521
noninsulin-dependent diabetes mellitus (NIDDM), 212
non-small cell lung cancer, 121B
nonspecific defences, 100, 101F
nonsteroidal anti-inflammatory drugs (NSAIDs), 146, 147
noradrenaline, 193, 422
norepinephrine *see* noradrenaline
normal ranges, 10
normal value, 10
normoblasts, 385
normochromic and normocytic anaemia, 299
normocytic anaemias, 396–397, 397T, 398
nose jobs, 174B
nosocomial infections, 55
nosocomial pneumonia, 78
notifiable diseases, 6
NP complex (NPC), 492, 493–494, 493F, 495
NPC *see* NP complex
NPs *see* nuclear pores
NR *see* nuclear receptor
NSAIDs *see* nonsteroidal anti-inflammatory drugs
NSCLC *see* non-small cell lung cancer
NTDs *see* neglected tropical diseases
NTP *see* National Toxicology Program
nuclear envelope, 492, 493F, 494
nuclear factors of activated T cells (NFATs), 183
nuclear import receptors, 493F
nuclear lamins, 519–520B
nuclear locating signal (NLS), 198–199, 199F, 472B, 472F, 493, 495
nuclear pores (NPs), 492, 493F
nuclear radiation *see* radiation
nuclear receptor (NR), 471B
nuclear skeleton, 491–492
nucleases, 319
nuclei and nucleocytoplasmic transport, 491–495
nuclei staining, 534–536
nucleic acids, 291
  Oxford Nanopore sequencing method, 517F
  see also amino acids
nucleocytoplasmic transport, 493–495
  see also nuclear locating signal
nucleoplasm, 492, 493F
nucleoporin 98 (*NUP98*) genes, 495
nucleoporins, 492, 493, 493F, 494–495
nucleosides, 329, 330F
nucleotide binding domains (NBDs), 483
nucleotide excision repair (NER) pathway, 528–529
nucleotides, 329, 330F
  phosphate intake, 291
nullisomy, 473
nutrients
  and energy, 280
  overview of general roles, 279F
  see also diet; digestion
nutrition, 279
  ageing, 569
  diet and, 279–293
    carbohydrates, 280F, 281
    lipids, 281–283
    minerals and trace elements, 290–293
    nutrients and energy, 280T
    proteins, 283
    vitamins, 284–290
    water, 283
  nutritional disorders, 5–6, 293–302
    anorexia nervosa, 294, 296–297
    bulimia nervosa, 296–297
    general management, 307–309
    investigation and laboratory testing, 303–307
    minerals and trace elements, 301–302
    obesity, 294–297
    PEM, 293–294
    of vitamin nutrition, 297–301

O

OADs *see* organic acid disorders
OAR *see* organs at risk
obesity, 6, 212, 294–297, 358B, 433, 447, 526, 528
  assessment, 295B
  clinical complications, 294, 295T
  general management, 308–309
  incidence in England, 296, 297F
  investigation and laboratory testing, 304
  leptin, 296B
  schematic/characteristic shape, 294, 295F
occupational therapy polyploidy, 476
OCT *see* optical coherence tomography
octreotide, 206F
oedema, 170, 246–247, 294, 300, 311C, 343, 348, 372, 436–437
oesophageal ulcers, 340
oesophagus, disorders, 337–341
oestradiol, 198F, 227F, 228, 542–543
oestrogen, 121B, 228, 229B, 230, 517M, 528, 542–543
  cytochrome P-450 aromatase, 542, 551F
oestrogen HRT, 549
oestrogen receptors (ERs), 542–543, 543T, 549–550, 551M
oethinyloestradiol, 229B
offensive virulence factors, 36–41
Okinawa, Japan, 283
olaparib, 546
oligomenorrhoea, 230

oligonucleotides, 506, 506M, 516B
oligosaccharides, 152, 162, 163F, 167
oliguric phase of AKI, 243
olorofim, 92M
ω carbon atoms, 281–283
see also fatty acids
ω-3 fatty acids, 282F, 527, 532B
Onchocerca volvulus, 55
onchocerciasis (river blindness), 49B, 55
oncogenes, 517, 518T
oncogenic fusion proteins, 495
oncology, 514
artificial intelligence, 548B
see also cancer
oncoprotein BCR-ABL, 522F, 523F, 562
oocytes, 453, 474, 476
see also menstrual cycle
oogenesis, 225–226
see also menstrual cycle
ophthalmology tests, 562B
opisthotonos, 299
opium poppy (*Papaver somniferous*), 358F, 358M
opportunistic infections, 2–3
opportunistic pathogens, 24
opsin proteins, 289
opsonins, 139
opsonisation, 106, 108, 112T, 140, 185
optical coherence tomography (OCT), 562–563B
oral cancer
alcohol and cancer, 528
Burkitt lymphoma, 529
cigarettes and smoking, 525–526B
oral contraceptive pill, 229B
oral disease, 82
see also mouth
oral dissolution therapy, 350F
oral feeding, 307–308, 308F
oral glucose tolerance tests (OGTT), 206, 235C
oral herpes, 337
oral hygiene, 337, 337M
oral infections, 58–59
oral rehydration, 308, 309F
oral supplementation, 308
Orange-G (OG-6), 535
orbital rhabdomyosarcoma, 542
orchiectomy, 558
organ donor cards, 180F
organic acid disorders (OADs), 270
organs at risk (OAR), 548B
orlistat, 308–309, 309F
ornithine decarboxylase, 88F
orthoherpesviridae, 529T
orthopaedic appliances, 491, 506
osmolality, 245B
osmotic diuresis, 213
osteoarthritis, 294, 295T, 579F
osteoclasts, 251
osteomalacia, 300
osteoporosis, 300–301
osteosarcomas, 495
otitis media, 55–56, 73
ototoxicity, 77B
ovarian cancer, 521, 528
anti-angiogenic drugs, 547
antimicrotubular agents, 540
platinum-based drugs, 541
topoisomerase inhibitors, 540
ovarian failure, 230–231
ovaries, 224F, 225F, 474
PCOS, 230, 231
overactive thyroid, 433
overhydrated, 244T
overweight, 235C, 310C
see also body mass index; obesity
ovum, 228F
see also oocytes; oogenesis
oxacillin, 124B
oxalate, 411
oxaliplatin, 541F
Oxford Nanopore sequencing method, 517B
oxidases, 138F
oxidation, 265, 284–285, 293, 501F
adrenoleukodystrophy, 499–500
endoplasmic reticulum, disorders of, 487
G6PD, 409
Refsum disease, 500
oxidation reactions, 363B
oxidative decarboxylation, 284
oxidative phosphorylation, 76, 335–336, 367
oxidative stress, 409–410, 519B, 524, 527, 528
oxidised glutathione (GSSG), 572–573, 572F
oxygen, 422, 423F
acute heart failure, 434
angina, 438–439
binding curves, 403
pressure (PO_2), 271–272
ROS, 525, 572, 572M
stroke, treatment, 446
see also cardiovascular system; dioxygen
oxygenase, 333
oxyhaemoglobin (HbO_8), 264–265, 264F
oxytetracycline, 78
oxytocin, 200F, 202F, 204

P

P waves, 428–429
p53 nonhistone protein, 547
p110α catalytic subunit, 517M
P-450 systems, 91F, 362F, 363B, 363F, 365, 369–370, 370F, 373, 408M
p450s
PABA *see p*-aminobenzoic acid
PacBio, 517B
pacemakers, 422–423, 431
Pacific yew tree (*Taxus brevifolia*), 456F, 456M, 540
packed cell volume (PCV), 396, 397, 397B, 397T
paclitaxel, 540
PAGE *see* polyacylamide gel electrophoresis
PAHs *see* polycyclic aromatic hydrocarbons
PAHX *see* α phytanoyl CoA hydroxylase
palliative care, 517B, 535, 536
palpation, 416C
palpitations, 396, 431
p-aminobenzoic acid (PABA), 348–349
PAMP *see* pathogen-associated molecular patterns
Pan-Cancer Analysis of Whole Genomes (PCAWG), 516
pancreas, 315–318, 326–328, 326–327B
blood glucose regulation, 211–212
cystic fibrosis, 486
disorders of the exocrine pancreas, 348–349
ethanol toxicity, 371
see also islets of Langerhans
pancreatic cancer, 348
alcohol and cancer, 528
cigarettes and smoking, 525–526B
diarylheptanoids, curcumin and its anticancer properties, 527
nucleocytoplasmic transport, 495
and obesity, 528
pancreatic duct cells, 326–327B, 327F
pancreatic esterases, 328
pancreatic juice, 318, 325–328, 336
pancreatic lipase, 327–328, 327F, 328M
see also pancreatic juice
pancreatic polypeptide, 336
pandemics, 30–32
COVID-19, 33F, 69, 74, 74F, 103M, 116B, 116F, 121B
see also plague
Paneth cells, 325
panitumumab, 553
pantothenic acid, 286
see also vitamin B_5
PAP *see* prostatic acid phosphatase
Papanicolaou (Pap) stain, 535
papaverine, 358F, 358M
papillary sarcomas, 495
papillomaviridae, 529T
PAPS *see* 3′-phosphoadenosyl 5′-phosphosulphate
papules, 300
paracentric inversions, 471F

paracetamol (acetaminophen), 349, 364–367, 380C
poisoning
clinical features, 365, 366T
laboratory testing, 366–367
management, 367
toxic doses, 365
paraesthesiae, 310C
parasites and pathogens, endoplasmic reticulum, disorders of, 487
Paramyxovirus (mumps), 415
paraparesis, 500
paraplegia, 399
paraquat, 374–375
parasites *see* pathogens and parasites
parathyroid hormone (PTH), 249, 252F, 253F
Ca^{2+} homeostasis disorders, 249–251
Mg^{2+} homeostasis disorders, 255
parental feeding, 308
parenteral hypercaloric feeding, 532B
parietal cells, 326B, 326F
Parkinson disease (PD), 5, 302, 520B, 575, 580, 581M
PARP *see* poly(ADP-ribose) polymerase
Patau syndrome, 473–474, 474F
pathogen-associated molecular patterns (PAMPs), 105, 125
pathogenecity, 36
pathogenesis, 3, 4F, 23
pathogens and parasites, 1F, 23–52, 53F, 85, 342
activation of T_H cells, 123
anthelmintic drugs, 92–94
antibacterial drugs, 76
antifungal drugs, 89–92
anthelmintic drugs, 92–94
antiprotozoal drugs, 86–89
arthropods and vertebrates, 48–50
bacteria, 34–41
controlling the spread of, 69
course of an infection, 25
and culture, 70
cytokine release syndrome, 103M
deaths from infections, 25–26
drug receptors, 359
Duffy blood group system, 167
eosinophils, 109
fungi, 45–46, 89–92, 125
helminths, 46–48, 92–94
helper T lymphocyte activation, 123
humoral immunity, activation of B lymphocytes, 120
infections of the CNS, 56
infections of the GIT, 58–60
infections of the urogenital system, 60, 61F
infectious diseases, 24–25
inhibitors of DNA replication/transcription, 76
lactose intolerance, 342
nonspecific defences, 100
prions, 26
protozoa, 41–45, 86–88
types of immune defence, 100
vaccines, 67–68, 67F, 69M
see also fungi; helminths; protozoa
pathology, 1
pathology laboratories, 8–10
pattern recognition receptor (PRR), 105, 125
Pauling, Linus, 403B
PAX3-FOXO1 hybrid protein, 471
PBMCs *see* peripheral blood mononuclear cells
PBSCT *see* peripheral blood stem cell transplant
PCAWG *see* Pan-Cancer Analysis of Whole Genomes
PCFT *see* H^+-coupled folate transporter
PCOS *see* polycystic ovary syndrome
PCP *see Pneumocystis carinii* pneumonia
PCR *see* polymerase chain reaction
PCV *see* packed cell volume
PD *see* Parkinson disease
PDE *see* phosphodiersterase; phosphodiesterase
PDE5 inhibitors (PDEIs), 233B, 233F
PDEIs *see* PDE5 inhibitors
PDGF *see* platelet-derived growth factor
PD-L1 *see* programmed death-ligand 1
peak expiratory flow (PEF), 149
peak flow meters, 150B, 151F
pefloxacin, 76
PEG *see* polyethylene glycol
pellagra, 299, 371
PEM *see* protein–energy malnutrition
pembrolizumab, 121B, 544
penicillins, 80, 80M, 152
Penicillium spp., 80F, 80M
pentamidine, 86
pentobarbitone, 373F
pentolinium, 223B, 224F
pentose phosphate pathway (PPP), 87–88, 87F, 408–409
pepsin, 323F, 339, 535
peptic ulcers, 251, 340–341
Zollinger–Ellison syndrome, 341
peptide mapping, 401B, 401F, 403B, 403F
peptide YY, 336
peptides, 192
absorption, 328, 329F
see also amino acids
peptidyltransferase, 79
perforins, 122, 123F
pericardial disease, 445
pericardium, 422
pericentric inversions, 471F
periodic fever syndrome, 115, 135B
periodontitis, 337
peripheral arterial disease, 445
peripheral blood mononuclear cells (PBMCs), 179–180
peripheral blood stem cell transplant (PBSCT), 132–133B, 185, 186, 187
peristalsis, 322
peritoneal cancer, 547
peritoneum, 315, 318–319
peritonitis, 320B
periwinkle (*Vina rosea*), 540F
pernicious anaemias, 300, 323M, 340, 349, 399
peroxidases, 284M, 290F, 293, 307F
peroxidation, 375F
peroxisomal disorders, 499–502
adrenoleukodystrophy, 499–500
Refsum disease, 500–501
rhizomelic chondrodysplasia punctata, 501–502
Zellweger spectrum disorder, 502
peroxisomal targeting signal (PTS), 499, 501T
personalised medicine, 535, 536M
pertussis exotoxins, 39–40
PET *see* positron emission tomography
petechiae, 415
petechial haemorrhages, 300
PEX gene mutations, 499, 500, 502B
Peyer patches, 59F, 60, 66
PGF *see* placental growth factor
pH, 318, 325, 326
aspirin toxicity, 367, 369
barbiturates, 373
definitions, 266B
erythrocytes, 385
lysosomal storage diseases, 488, 489F
SCA and malaria, 405B
see also H^+
PHA *see* phytohaemagglutinin
phagocytes, 129
apoptosis and cancer, 520B
see also macrophages
phagocytic defects, 138–139
phagocytic system, mononuclear, 317
phagocytosis, 36–37, 58, 93, 108, 110, 254, 257, 488
complement, 103, 106F
lysosomal storage diseases, 488
phagosomes, 488
phagotherapy, 84–85B
pharmacodynamics, 360
pharmacokinetics, 360
pharmacology, 357–359
pharmacotherapeutics, 359
pharyngeal cancer, 528
phenobarbitone, 299–300, 373F
phenothiazines, 202F

phenotypes, 165–166, 166T, 458, 466–467, 502
phenotypic resistance, 61B
phenoxybenzamine, 223B, 224F
phenylalanine, 11–12, 344B, 464–465, 464F, 464T, 479C
phenylalanine hydroxylase, 5
phenylketonuria (PKU), 5, 454, 462T, 464–465, 486
phenytoin, 299–300
pheochromocytomas, 223–224B, 436–437
Philadelphia (Ph) chromosome, 522F, 523F, 533T, 560, 562
phosphate (P_i), 213, 254–255, 267, 290–291, 291T, 307
phosphatidylcholines, 38F
phosphatidylinositol 4,5-bisphosphate (PIP_2), 517M
phosphatidylinositol 3,4,5-triphosphate (PIP_3), 517M
phosphatidylinositol-3-kinase (PI3K), 517M, 518T, 522, 545–546
  angiogenesis, 546
phosphatidylserine, 392–393, 393F, 519–520B
3′-phosphoadenosyl 5′-phosphosulphate (PAPS), 362, 364
phosphodiesterase (PDE), 196, 232–233B
phosphoenolpyruvate, 80F
phosphoenolpyruvate carboxykinase, 293F
phospholipase A2, 76
phospholipases, 38F, 76
phospholipids, 291, 392–393, 393F, 411–412, 439–441
phosphorylases, 367, 368F
phosphorylation, 495, 517M, 521
phototherapy, 170, 195B, 349–350, 349M
PHYH gene mutations, 500
phylloquinone (K_1), 290F, 394B, 435B
physical dependence, barbiturates, 373
phytanic acid (phytanate), 500, 501F
phytate, 302, 303F
phytohaemagglutinin (PHA), 131B, 131T, 132F, 466
PI3K *see* phosphatidylinositol-3-kinase
PI3K/AKT kinase signalling pathway, 545–546
PIDs *see* primary immunodeficiency diseases
pig experiments, 175–176B
pigments, 574F
PIK3CA mutations, 517M, 533T, 536
pill, the, 229B
pilocarpine, 485–486, 485F
Pincus, Gregory, 229B
pineal gland, 194–195B, 194F
pinworms, 47–48, 92T
PIP_2 *see* phosphatidylinositol 4,5-bisphosphate
PIP_3 *see* phosphatidylinositol 3,4,5-triphosphate
piperaquine, 86
pituitary gland, 200–204, 208, 229B, 230
  function disorders, 202–204
  water disorders, 244–245
  see also anterior pituitary; posterior pituitary
pituitary gonadotropin, 543
PKG *see* protein kinase G
placental growth factor (PGF), 546
plague, 34–35B, 48
Plague of Justinian, 35B
plaques, 4, 95C, 337, 438, 575, 576F
  atherosclerosis, 300, 438, 439–440, 444
  see also blood clotting
plasma, 160F, 383T
  adrenoleukodystrophy, 500
  blood clotting, 389–395
  type II hypersensitivities, 152, 153F
  uric acid concentration, 256F, 256M, 257, 301, 302
  see also blood
plasma cell tumours, 126C
plasma cells, 120F
  monoclonal antibodies, 120–121B
plasma membrane-related disorders, 29T, 483–486
  ABO blood groups, 162, 163F
  apoptosis and cancer, 520B
  cholesterol, 439–440
  cytoskeletal disorders, 503–504, 504F
  G6PD, 409
  graft rejection, 175
  haemoglobinopathies, 401–402
  immunosuppression, 185F
  mitosis/meiosis, 456
plasma proteins, 103–107
  biological variation, 17
  Lewis blood group system, 166–167
  see also complement
plasma thromboplastin antecedent (PTA), 390T, 392, 414M
plasmapheresis, 173M
plasmin, 395F, 439M
plasminogen, 349M, 395
Plasmodium falciparum, 409M
Plasmodium spp., 36B, 41, 41–42B, 125, 405B
  see also malaria
Plasmodium vivax, 167
platelet glycoproteins IIIA and IIb, 439
platelet-derived growth factor (PDGF), 392F, 525B
platelet-derived growth factor receptor α (PDGFRα), 547
plateletpheresis, 173M
platelets, 136, 160F, 383T, 384F, 386, 436B
  aggregation, angina, 439
  concentrates, 387M
  electron micrograph, 386, 387F
  haematopoietic system, 395–396
  sickle cell anaemia, 405
  thrombocytopenia, 414–415
  thrombocytosis, 415
  Wiskott–Aldrich syndrome, 136–137
  see also blood clotting
platinum-based drugs, 541
platyhelminths *see* flatworms
PMNs *see* polymorphonuclear leukocytes
Pneumococcus spp., 68B, 137, 404
Pneumocystis carinii, 58, 129
Pneumocystis carinii pneumonia (PCP), 62–63B
pneumonia, 32, 55, 56M, 58, 67, 68B, 78, 129, 154C, 273B, 505, 555
PNP *see* purine nucleoside phosphorylase
POCT *see* point-of-care testing
podophyllotoxin, 540
point-of-care devices, 429
point-of-care testing (POCT), 14, 14M
poisoning, 349, 361, 361M
  biochemical investigations, 364T
  common poisons, 364–379
    aspirin, 367–369
    barbiturates, 373
    carbon monoxide, 373–374
    ethanol, 369–371
    ethylene glycol, 372
    metals, 375–379
    methanol, 372
    paracetamol, 364–367
    paraquat, 374–375
  symptomatic, 364
  see also food poisoning; toxicity
poke weed mitogen (PWM), 131T
pokilocytosis, 398
polar bear (*Ursus maritimus*), 301F
polio, SCID, 131
poliomyelitis viruses, 56
poliovirus, 68–69M
pollen, 151F
polyacylamide gel electrophoresis (PAGE), 485
3′polyadenylation, 407, 407M
poly(ADP-ribose) polymerase (PARP), 546
polyallelic genes, 177
polychlorinated biphenols, 524T
polyclonal activators, 131B, 131T
polycyclic aromatic hydrocarbons (PAHs), 524T, 555
polycystic kidney disease, 244
polycystic ovary syndrome (PCOS), 230, 231
polycythaemia vera, 257–258
polydipsia, 203, 212, 248
polyenes, 89, 90–91
  see also amphotericin B
polyethylene glycol (PEG), 120B

polymerase chain reaction (PCR), 67, 71–72, 168, 180F, 401B, 407–411, 478, 485, 535
and Hubert H. Humphrey, 73M
polymerases, 54B, 62B, 77F, 496
polymerisation, 537, 539–540
polymixins, 82–83, 83F
polymorphonuclear leukocytes (PMNs), 108, 109, 109T, 386
polyomaviridae, 529T
polyomyelids, 27, 27F
polypeptides, 492, 496
ageing, abnormal proteins and the error-catastrophe theory, 574
genetics and DNA, 454–455
see also amino acids; proteins
polyploidy, 476–477
polyprenly pathway, 89, 90F, 135B, 135F
polysaccharides, 35–36, 36F, 68–69
anticoagulant therapy, 435B
polyuria, 203, 212, 245, 248
pork tapeworm, 92F
porphyrias, 461
porphyrins, 461
portals of entry/exit, 25
positron emission tomography (PET), 7, 541, 555, 577–578, 578F
posterior pituitary, 246–247
posthepatic hyperbilirubinaemia, 349–350
posthepatic jaundice, 349
postmenopausal women, 542–543, 546, 549–550
postnecrotic cirrhosis, 353
postrenal injury, 243
post-transplant support, 187
potassium (K^+), 89, 213, 222, 245–246, 248–249, 270, 328, 372, 429
homeostasis disorders, 248–249
hypokalaemia, 249
potassium serum concentrations, 307
Potts, Percivall, 524
PPD *see* purified protein derivative
PPP *see* pentose phosphate pathway
PR intervals, 428
Prader–Willi syndrome, 470
praziquantel, 93F
pRB nuclear phosphoprotein, 518, 520, 521M, 562B
precision, 15
predisposing factors, 2
prednisolone, 181T, 182F, 506, 532B
pregnant women, 5, 228
adrenoleukodystrophy, 500
amenorrhoea, 230–231
antibody deficiencies, 133, 136
anticoagulant therapy, 436B
Bart hydrops, 406T, 407–408
Chédiak–Higashi syndrome, 138
cholelithiasis/cholestasis, 351B
DiGeorge anomaly, 136
fetal alcohol syndrome, 136
folic acid deficiency, 299–300
Graves' disease, 143
HbF, 388, 389M, 400B, 401B, 401F, 404–405, 407, 408, 416C
HPFH, 389M
human milk, lactose intolerance, 341–342
immunological tolerance, 141
iron metabolism, 385B
Kleihauer testing, 171F
LSDs, diagnosis and treatment, 491
NALD, 499–500, 502
normocytic anaemias, 398
thalassaemias, 407–408
thrombocytopenia, 415
thrombocytosis, 415
umbilical cord stem cell transplantation, 187
vitamin A deficiency, 300
vitamin K deficiency, 300–301
WAS gene mutations, 137
X-irradiation *in utero*, 528
see also adult/maternal Hb; neonates
prehepatic hyperbilirubinaemia, 349
prehepatic jaundice, 349
prekallikreins, 392
premalignant cancer, 2
premature cell lysis, 409
premutations, 467
prepoinsulin/active insulin conversions, 211F
prerenal injury, 243
prevalence of disease, 6
prevention, control and treatment of diseases, 53–98
controlling the spread of pathogens and parasites, 69
infections of the gastrointestinal tract, 58–60
infections of the respiratory system, 57–58
infections of sense organs and the CNS, 55–56
infections of the skin, 53–55
infections of the urogenital system, 60–65
investigating infectious diseases, 69–74
preventing infectious diseases, 53, 67–69
sepsis and systemic infections, 65–67
see also vaccines and vaccinations
prilocaine, 395
primaquine, 78, 409
primary acquired lactose intolerance, 341–342
primary disease, 3
primary hyperparathyroidism, 251
primary hypogonadism, 230
Primary Immunodeficiency Diseases Expert Committee, 130
primary immunodeficiency diseases (PIDs), 125, 129, 130–140
antibody deficiencies, 130, 133–136
Chédiak–Higashi syndrome, 137–138
complement deficiencies, 140
DiGeorge anomaly, 130, 136
phagocytic defects, 138–139
SCID, 130, 131–133
Wiskott–Aldrich syndrome, 136–137
X-linked congenital neutropenia, 136–137
X-linked thrombocytopenia, 136–137
primary resistance, 61B
primary thrombocytosis, 415
primary tumours, 514
primidone, 299–300
prion diseases, 462–463B
see also spongiform encephalopathies
prions, 2M, 26
proapoptotic proteins, 519B
procaspase, 519B
processed food, 526–527
progeria, 5, 581–582B, 583C
progession (tumourigenesis), 516F, 524
progesterone, 121B, 221F, 228, 229B, 447, 517M
progesterone HRT, 549
prognosis, 5
programmed ageing, 574–575
programmed cell death protein-1 (PD-1), 121B, 544
programmed death-ligand 1 (PD-L1), 544
proinsulin, 211F
prolactin, 202–203, 202F, 203T, 230–231, 230F
prolactinoma, 202–203
proline, 344B
promotion (tumourigenesis), 516F, 524
promyelocytes, 472B
properdin, 140
prophase, 456F, 458
prophylactic bilateral mastecomy, 521, 522B
prophylactic treatment, 442
prophylaxis, HDN, 171
Propionibacterium spp., 54, 55F
propionic acids, 270F
propranolol, 431F, 445–446
prostacyclins, 281, 436B
prostaglandins, 93, 148–149, 196, 239–240, 251, 281, 340–341
prostate cancer, 19C, 514, 528, 533T, 556–559
antimicrotubular agents, 540
carotenoids, lycopene, 527
Gleason scores, 557–558, 559T

hormone therapy, 542
prostate-specific antigen, 533T, 557, 558, 563C
signs, symptoms and diagnosis, 557–558
TNM system, 557–558
cancer treatment, 558
treatment, 558–559
prostatectomy, 558
prostate-specific antigen (PSA), 533T, 557, 558, 563C
prostatic acid phosphatase (PAP), 544
prostatis, 60, 61F
prostitution, 172–173
protease caspase 1, 102–103
protease inhibitors, 64B, 390–391
proteases, 109T
proteasome inhibitors, 520B
proteasomes, 518
protein conformational diseases, 463B
protein kinase A, 39–40, 40F
protein kinase G (PKG), 39, 232–233B
proteinase K, 535
protein–energy malnutrition (PEM), 293–294, 297–298, 302
general management, 308
investigation and laboratory testing, 303–304
proteins, 9F, 280F, 283
ABC, 441M, 483, 484F, 499–500
adrenoleukodystrophy, 499–500
ageing, abnormal proteins and the error-catastrophe theory, 574
ageing, glycated proteins, 571, 573
biological variation, 17
blood clotting, 389–395
blood disorders, erythrocytes, 385
cachexis, 531–532B
chronic myelogenous leukaemia, 522F, 523F, 562
coiled-coils, 488M
C-reactive protein, 70, 107B, 432B, 439
cytoskeletal disorders, 503–504
and diet, 283
embryonic/fetal haemoglobin, 388F
endoplasmic reticulum, disorders of, 486–487
epigenetic regulators, targeted therapy, 547
G protein-coupled receptors, 196, 197F
Golgi apparatus, disorders of, 487–488
graft rejection, 175
heat shock proteins, 198–199, 198M, 199F
helper T lymphocyte activation, 123
hereditary elliptocytosis/spherocytosis, 506, 507T
hormone classification, 192
humoral immunity, activation of T lymphocytes, 122
iron absorption, 333
lysosomal storage diseases, 488, 489T
mitochondrial membrane, 495T
muscular dystrophies, 504–506
nucleocytoplasmic transport, 493–494, 495
offensive virulence factors, 36–41
plasma membrane-related disorders, 483, 484F
RCDP, 501T
Rhesus (Rh) systems, 165
serum protein separation, 112F
spike proteins, 27, 100F
topoisomerase inhibitors, 540
vaccines, 67–69
viruses, 27
WAS-related disorders, 136–137
zymogens, 325M, 327T, 390–391, 439M
see also amino acids; complement; enzymes; glycoproteins; immunosuppressants/immunosuppression; lipoproteins; marker proteins; membrane proteins
proteinuria, 351
proteolysis-inducing factor, 531–532B
proteolytic enzymes, 326–327
Proteus spp., 78
prothrombin, 392–393, 393F, 394B, 411–412, 435B
see also blood clotting
protomer proteins, 27
proton beam therapy, 542, 563B
proton pump, 326
protonated haemoglobin (H^+Hb), 264
proto-oncogenes, 517, 518T, 529
protozoa, 41–45, 85
amoebiasis, 43
antiprotozoal drugs, 85–89
giardisis, 43–44
immune response, 125
inhibitors of DNA replication/transcription, 76
leishmaniasis, 43, 44
toxoplasmosis, 43, 44–45
trichomoniasis, 43, 45
trypanomiases, 43
provitamin A carotenoids, 287–288, 288F
proximal tubules, 240, 241F
PRR *see* pattern recognition receptor
PSA *see* prostate-specific antigen; prostate-specific antigen
P-selectins, 405
PSEN1 gene mutations, 577
pseudoautosomal region (PAR), 460F
pseudohypertrophic muscular dystrophy *see* Duchenne muscular dystrophy
pseudohypoparathyroidism, 250
Pseudomonas aeruginosa, 24F, 35–36, 54–55, 60, 78, 80, 83, 85B, 140
eye infections, 55
Pseudonomas aeruginosa, complement deficiencies, 140
pseudoprecocious puberty, 221–222
psoriatic arthritis, 121B
psychogenic diseases, 6
PTH-related protein (PTHrp), 251
PTHrp *see* PTH-related protein
ptosis, 144
PTS *see* peroxisomal targeting signal
puberty, 54
pubic louse (*Phthirus pubis*), 48F, 65
pulmonary artery, 424
pulmonary circulations, 422
pulmonary fibrosis, 374
pulmonary oedema, 172, 434, 436–437
pulse, 426
PUMA (p53 unregulated modulator of apoptosis), 519–520B
purified protein derivative (PPD), 153–154
purine, 182F, 258
purine analogues, 538–539
purine nucleoside phosphorylase (PNP), 132F
purines, urate balance disorders, 256F
Purkinje fibers, 422–423
purpura, 415
putrescine, 88
Pwo/*Pfu* polymerases, 73
pyelonephritis, 60, 61F
pyloric sphincter, 323
pyogenic bacteria, 133–134
pyrantel pamoate, 93
pyrazinamide, 75M, 82, 83, 83F, 85
pyridoxal, 286, 287F
see also vitamin B_6
pyridoxamine, 286, 287F
see also vitamin B_6
pyridoxine, 286, 287F
see also vitamin B_6
pyrimidine analogues, 538–539, 561
pyruvate, 335–336
pyruvate accumulation, aspirin toxicity, 367
pyruvyltransferase, 80
Pythagoras, 411

Q

Q waves, 429
QRS complex, 428, 431–432
QRS/QT periods, 176
QT intervals, 428

quality control, 18–19
quinolones, 76–77
quinupristin, 75M, 78

R

R waves, 429
RA *see* rheumatoid arthritis
Rab11a G protein, 196F, 196M
rabies virus, 56
RAD51-DNA complex, 458F
radial immunodiffusion (RID), 116B, 116F
radiation
  and cancer, 519B, 521, 528–529
  ionizing, 536
  UV, 289F, 521, 528–529, 571B
  see also ionizing radiation
radiation therapy (radiotherapy), 536, 541–542
  acute myeloid (myeloblastic) leukaemia, 561
  breast cancer, 551–552
  lung cancer, 556
  prostate cancer, 556
  retinoblastoma, 563B
radical formation, 138F, 374–375, 375F
  see also free radicals
radioallergosorbent test (RAST), 149
radioimmunoassay (RIA), 116B, 146, 533
radioisotopes, 542T
radon gas, 555
RAG1 and *RAG2* genes, 132
Ran G protein, 493–494
rapamycin, 184F
RARE *see* retinoic acid response elements
RAS/RAF/MAPK, PI3K/AKT/mTOR pathways, 522
Ras-associated protein, 488, 489T, 546
rash, 29F, 66–67
  urticaria, 152
  see also allergies
Rb *see* retinoblastoma
RB1 gene mutations, 518–520, 521M
  see also retinoblastoma
RCDP *see* rhizomelic chondrodysplasia punctata
RD *see* Refsum disease
RDA *see* recommended daily amounts
reactive oxygen species (ROS), 525, 572, 572M
Rebuck skin window test, 155C
receptor tyrosine kinases (RTKs), 197, 198F, 211
receptors
  activation of complement, 104–106
  acute promyelocytic leukaemia, 471–472B, 472F
  anthelmintic drugs, 93
  apoptosis and cancer, 520B
  basophils, 109
  chemokines, 103F
  complement and blood transfusions, 169T
  cytokines, 101F
  fat-soluble vitamins, 288–289
  GnRH, 543, 558–559
  intracellular receptors, 198–199
  neutrophils, 108
  nucleocytoplasmic transport, 493F
  oestrogen, 542–543, 543T, 549–550, 551M
  opsonisation and inflammation, 106F
  plasma membrane-related disorders, 483
  PRR, 105, 125
  RAR, 229B, 471–472B
  regulation of blood glucose, 211
  second messengers, 195–197, 198F
  T cell receptors, 118, 122–123F, 185F, 545F
    distinguishing between B and T lymphocytes, 117–118, 119B, 119F
    and immunoglobulin domains, 113F, 113M
    superantigens and toxic shock syndrome, 124B, 124F
  targeted therapy, 545–546
recessive alleles, 459, 460–461, 461T
recirculation, 117F
recommended daily amounts (RDA), 280
recreational drugs, 395
rectal cancer *see* colorectal cancer
red blood cells (RBCs) *see* erythrocytes
red meat in diet, 527
red-green colour blindness, 461T
reduced folate carrier (FOLT), 331
reduced glutathione (GSH), 572–573, 572F
5α-reductase, 226
reductases, 86, 87F, 332–333, 359F, 387F, 409, 441–442, 442F
  DHFR, 537, 538F
  glutathione, 86, 285, 304–305, 305F
reduction division, 456
reference nutrient intake (RNI), 280F
reference ranges, 10, 10–11B
reflux oesophagitis, 338
Refsum disease (RD), 500–501
Refsum, Sigvald, 500
regurgitation, 437–438
rehydration, 308, 309F
  see also dehydration
rejection of transplants, 45, 55, 159, 175–176B, 180–181
  see also grafts
relapse, 5
relatlimab, 544
remission, 5, 214B
Renaissance Italy, 174B
renal calculi, 243
renal cancer, 543
renal compensation, 269
renal function, 70
  renal failure, 12
  see also kidneys
renal function tests, 242–243
renal tubules, 372
renin, 222, 239–240, 434M
repair molecules, 546
replication, 82
  antifungal drugs, 90F
  bacteriophages and phagotherapy, 84B, 84F
  endoplasmic reticulum, disorders of, 487, 487T
  mitochondrial DNA, 496
reproductive hormones, 224–233
  disorders, 230–231
  fertilisation and pregnancy, 228, 229B
  infertility, 231–232
  menopause, 228
  oral contraceptive pill, 229B
  see also oestrogen; progesterone; testosterone
RER *see* rough endoplasmic reticulum
respiration, *see also* lungs
respiratory acid–base disorders, 268, 270–273
respiratory acidosis, 271–272
respiratory alkalosis, 272
respiratory compensation, 269
respiratory complex III, 499
respiratory disorder, 294, 295T
respiratory system
  adenoviruses, 1F
  infections of, 57–58
  see also lungs
resting membrane potential, 328
restriction endonuclease *Mst*II, 404F
restriction fragment length polymorphisms (RFLPs), 404, 404F, 478
restrictive cardiomyopathy, 437
reticulocytes, 385, 385B
reticulocytosis, 401–402
reticuloendothelial system, 317
retinal, 287–288
retinal-binding proteins, 300F
retinaldehyde, 287–288, 289
retinoblastoma (Rb), 518, 520, 521M, 542, 562–563B
retinoblastoma susceptibility gene (RB1), 518–520
retinoic acid, 192, 287–288
retinoic acid receptors (RAR), 229B, 471–472B
retinoic acid response elements (RARE), 471B
retinoid signalling pathways, 198–199

retinoid X receptors (RXR), 471B
retinol, 287–288
retinopathy, 4, 213, 448
retroviridae, 529T
retroviruses, 529–530
see also human immunodeficiency virus
reverse grouping/typing, 168, 168M
reverse transcriptase (RT), 62–64B, 529–530
reverse transcription PCR (RT-PCR), 73, 74
reverse triiodothyronine (RT_3), 209F
reversion, 68
rezafungin, 92M
RF *see* rheumatoid factor
RFLPs *see* restriction fragment length polymorphisms
RFVTs *see* riboflavin transporters
rhabdomyosarcoma
antimicrotubular agents, 540
proton beam therapy, 542
topoisomerase inhibitors, 540
RHCE gene, 165
RHD gene, 165
Rhesus (Rh) systems, 162, 165–166, 168F, 170, 171, 188C, 349
rheumatic fever, 432, 432B
rheumatoid arthritis (RA), 121B, 140, 144–146, 153, 283, 367, 520B
rheumatoid factor (RF), 144–146, 145F, 153
rhinitis, 148, 148M, 148T
see also hay fever
rhinoplasty, 174B, 174F
rhinovirus, 57F
rhizomelic chondrodysplasia punctata (RCDP), 501–502
rhizomes, 527F
RIA *see* radioimmunoassay
riboflavin, 285F
see also vitamin B_2
riboflavin transporters (RFVTs), 331
ribonucleic acid (RNA), 319, 329, 454–455, 454F, 467, 467M, 529–530
acute myeloid (myeloblastic) leukaemia, 561
antifungal drugs, 89–92
Baltimore classes, 28T
CRISPR, 545, 547–548
folic acid antagonists, 537
gene therapy, 547–548
haemoglobinopathies, 401F
molecular biology of cancer, 515
NGS methods, 517B
progeria, 581–582B
purine/pyrimidine analogues, 538
TERT polymerase, 516B
see also messenger RNA; microRNAs; mRNA; transfer RNA
ribonucleoprotein (RNP), 548
ribonucleosides, 64B, 64F
ribosomal RNA (rRNA), 40, 79F, 496, 496F
ribosomes, 77–79, 454–455, 487, 491–492
ageing, abnormal proteins and the error-catastrophe theory, 574
mitochondrial DNA, 496
ribosylation, 39–40, 40F
ribosyltransferases, 39–40, 40F
rickets, 300F
Rickettsia spp., 36–37
Rickettsia typhi, 36–37
rickettsias, 124
RID *see* radial immunodiffusion
rifampicin, 75M, 77F, 435–436B
rifampin, 77, 83, 85, 415
rifamycins, 76, 77
ringworm, 125
risk factors, 2, 561
breast cancer, 549
colorectal cancer, 552–554
lung cancer, 555
rivaroxaban, 394B
river blindness, 49B, 55
RNA polymerase, 62B, 293
RNI *see* reference nutrient intake
RNP *see* ribonucleoprotein
Robertsonian translocation, 472F, 475F
Roche pyrosequencing, 517B
rolling hiatus hernias, 336–338, 338F
romiplostim, 415
Röntgen, Wilhelm, 7M
root (teeth), 320–321B, 321F
ROS *see* reactive oxygen species
Rose–Waaler test, 145F
rotational vertigo, 77B
rotaviruses, 27F, 59–60, 59F, 60, 68B, 70
rough endoplasmic reticulum (RER), 486F, 487, 488, 492F
roundworms, 60, 92T
Rous, F. Peyton, 529
Rous sarcoma virus, 529
RpsE, 79
RT *see* reverse transcriptase
RT_3 *see* reverse triiodothyronine
RT-PCR *see* reverse transcription PCR
rubella, 5, 415
rubella virions, 111F
Ruos, F. Peyton, 529
Russia, Chernobyl accident, 528

S

S adenosylmethionine (SAM), 468B, 468F
S phase of the cell cycle, 455F, 456M, 518, 520
acute myeloid (myeloblastic) leukaemia, 561
chemotherapy, 538, 539F, 540
Saccharomyces cerevisiae, 45F, 46, 68–69, 91F
complement deficiencies, 140
safe intake, 280
SAIDs *see* steroidal anti-inflammatory drugs
salbutamol, 151
salicylate, 269 *see* aspirin
saliva, 320, 321F
Salmonella spp., 404
Salmonella typhi, 66F, 79
Salmonella typhimurium, 66M, 526, 526F
Salmonellae, 36–37
salt baths, 348
salt methylglutaconate, 270F
SAM *see S* adenosylmethionine
sand flies (*Phlebotomus* spp.), 44, 48F
Sanger, Fred, 517B
Sanger, Frederick, 213M
saphenous system, 446–447, 446F, 447F
sarcoglycans, 504T
sarcoidosis, 95C, 251, 437, 437B
sarcolemma muscle fibers, 503F
sarcomas, 495, 514, 529–530, 552T
acute myeloid (myeloblastic) leukaemia, 561
antimicrotubular agents, 540
epigenetic regulators, targeted therapy, 547
immunotherapy, 543
platinum-based drugs, 541
topoisomerase inhibitors, 540
see also Kaposi sarcoma
sarcoplasm, 503F
see also cytoplasm
SARS *see* severe acute respiratory syndrome
scabies mite (*Sarcoptes scabiei*), 48F, 65
SCARB1 gene, 332
SCD *see* sickle cell disease
Schiff base, 573
Schilling test, 305–306
Schistosoma spp., 92F
schistosomes, 125
schistosomiasis, 46–47, 92F
SCID *see* severe combined immunodeficiency
SCLC *see* small cell lung carcinoma
sclerae, 416C
sclerosis, 438
Scopulariopsis brevicaulis, 45F
screening, 12–13M, 445M, 477, 479
blood, 172–173
cervical cancer, 531, 531M
cervical screening programmes, 531M
colorectal cancer, 553M
cystic fibrosis, 486
human papilloma viruses, 531

screening (*continued*)
PKU, 465, 465M
see also individual imaging techniques...
screening tests, 11–12, 12–13M
seasonal affective disorder (SAD), 195B
second messengers, 101, 195–197, 198F, 281
Second World War, 173, 525B
second-generation sequencing methods, 71
see also next-generation sequencing
secondary acquired lactose intolerance, 341–342
secondary diabetes mellitus, 212
secondary disease, 3
secondary immunodeficiency diseases (SIDs), 125, 129, 130T
secondary thrombocytosis, 415
secondary tumours, 3F, 514
second-generation sequencing, 517B
secretases, 575, 576F
secreted defensins, 57
secretin, 336, 392
secretors, 164
secretory component, 114–115
Sedlackova syndrome, 136M
selenium, 291T, 293, 302
selenium deficiency, 307F
semiconservative nature of DNA replication, 455F
senescence, 569
sense organs, infections of, 55–56
sensitivity, 15, 16–17B
sensitisation, 153–154, 185–186, 188C
allergic asthma, 149
allergic rhinitis or hay fever, 151
anaphylactic shock, 152B
blood and blood products for transfusion, 160T, 161
complement and blood transfusions, 169F, 169T
and HDN, 170, 170M, 171
immunoglobulin G, 113
Rhesus blood group, 165–166
SLA typing, 179
SRBCs, 145F
type I hypersensitivities, 148–149
sepsis and systemic infections, 65–67
septicaemia, 129
sequelae, 4F
sequence-specific oligonucleotide probe (SSOP), 180
sequence-specific primers (SSPs), 180F
sequencing methods, 71
see also next-generation sequencing
SER *see* smooth endoplasmic reticulum
serine proteases, 488
serine/threonine kinase, 9F
serological HLA typing, 179–180
serological tests, 70–71
serotonin, 109T, 344B, 532B
serotonin (5-hydroxytryptamine, or 5-HT), 194B, 344B, 391–392
serous fluid, 319
SERPINA1 gene, 272–273B
SERPING1 gene mutations, 140B
Serratia marcescens, 138, 543
serum, 383
albumin, 9F, 12F, 311C
folate, 305
protein separation, 112F
retinol-binding protein, 300F
see also plasma
SEs *see* spongiform encephalopathies
severe acute respiratory syndrome (SARS), 28T, 32, 33F, 36B, 74F, 100F
see also COVID-19
severe childhood autosomal recessive muscular dystrophy (SCARMD), 504T
severe combined immunodeficiency (SCID), 130, 131–133
sewage treatment plants, 23, 24F
sex chromsomes
aneuploidy, 475–476
see also XX chromsomes; XY chromosomes
sex hormone binding globulin (SHBG), 198F, 226, 227, 231
sex hormones, 141, 224–233
see also oestrogen; progesterone; reproductive hormones; testosterone
sex-linked genetic diseases, 450–451
sexually transmitted diseases (STDs), 55, 60, 62, 65, 350–352
investigating infectious diseases, 69
see also acquired immunodeficiency syndrome; human immunodeficiency virus
sGC *see* soluble guanylyl cyclase
sgRNA *see* single guide RNA
sheep red blood cells (SRBCs), 145F
Shiga toxins, 40F
Shigella dysenteriae, 24, 40F, 60
Shigellae, 36–37
shingles, 29F, 54
shingles vaccination, 67F
shock, 430, 445, 448
short bowel/gut syndrome, 345, 346–347
short hairpin RNA (shRNA), 405
shoshin beriberi, 298
Shprintzen–Goldberg syndrome, 136M
sibutramine, 309F
sickle cell anaemia (SCA), 385B, 402–405, 408, 409M, 416C, 454, 478, 506
diagnosing/treating, 404–405
inherited gene disorders, 461
and malaria, 405, 405B
sickle cell disease (SCD), 385B, 402, 404
see also sickle cell anaemia
sickle cell haemoglobin (HbS), 400B, 401B, 401F, 402, 403, 403B, 404F, 405B, 416C
SIDS *see* secondary immunodeficiency diseases
signal transducer and activator of transcription 3 (STAT3) protein, 135B
sildenafil (Viagra), 233F
silent mutations, 406
silica, 555
Simulium flies, 55
simvastatin, 359F, 441–442
single guide RNA (sgRNA), 548
single-molecule, real-time (SMRT) approaches, 517B
single-nucleotide polymorphisms, 521
single-nucleotide variations (SNVs), 516
single-stranded breaks (SSBs), 529
sinoatrial (SA) node, 422, 423F, 428, 429
sinonasal disease, 8F
sinusoids, 317
sipuleucel-T, 544
sirolimus (rapamycin), 184F
sirolimus–FKBP12 complex, 184
SIRS *see* systemic inflammatory response syndrome
site-1 protease (S1P), 488
skin, 53–55, 65–66, 99F, 125
ageing, 570, 570F
carcinogen-inducing tumours, 524F
chemotherapy, 537
eczema, 137, 148M, 334B, 464, 549
fat-soluble vitamins, 288–289
food allergies and intolerance, 149, 334–335B
Kaposi sarcoma, 129F
Langerhan cells, 123, 123M
leprosy, 49B, 49T, 55, 76, 96, 359
macupuscular skin rashes, 186M
methaemoglobinaemia, 395
and mucous membranes, 100M
paraquat, 374–375
radiation therapy, 541
rash, 29F, 66–67, 152
shingles, 29F
SLE, 129, 134, 140, 144, 146–147, 147T
and sunlight, 571B
thrombocytopenia, 415

thrombocytosis, 415
type I hypersensitivities, 148T, 149
type IV (delayed-type) hypersensitivies, 154
urticaria (skin rashes), 152
see also allergies
skip lesions, 345
SLC4A4 cotransporter, 266–267
SLC6A19 gene mutations, 344B
SLC14A1 gene, 167
SLC35C1 gene mutations, 139
SLE *see* systemic lupus erythematosus
sleeping sickness, 43, 56, 85, 86T, 88, 502B
sliding hiatus hernias, 336–338, 338F
small cell lung carcinoma (SCLC), 555
small intestine, 300, 308–309, 315–318, 316F
absorption of digested products, 328, 331
minerals and trace elements, 332–333
water, 333
bilirubin, 318
digestion, 319, 323, 325–328
disorders, 341–345
and homeostasis, 335–336
junction with the large intestine, 336F
peritoneum, 319
vitamin absorption, 331
vomiting and cyclic vomiting syndrome, 324B
see also pancreatic disorders
small lymphocytes, 109, 110, 117
smallpox, 30–31
smoking, 273B, 275C, 310C, 363B, 496–497, 525–526B, 555, 561, 564C
smooth endoplasmic reticulum (SER), 362F, 486F, 487
SMVT *see* Na^+-dependent multivitamin transporter
snail fever *see* schistosomiasis
SNVs *see* single-nucleotide variations
sodium chloride (table salt), 248, 291T
sodium dodecyl sulphate polyacrylamide gel elextrophoresis, 504F
sodium hydrogen carbonate ($NaHCO_3$), 213, 326, 327F, 369, 372, 373
sodium (Na^+), 39–40, 89, 217, 222, 245–246, 290, 291T, 320, 320F, 328, 344B
clotting pathways, 392F
diarrhoea, 346B
homeostasis disorders, 246–248
serum concentrations, 307
sodium/bicarbonate cotransporter (NBC 1), 266–267
SODs *see* superoxide dismutases
soluble guanylyl cyclase (sGC), 232–233B
somach infections, 59–60
somatic mutations, 459, 476, 536
somatomedins, 204
somatostatin, 204–205, 206F
somatotrophin, 204
see also growth hormone
sonograms, 7, 304
see also ultrasound
Spanish flu, 31
spasticity, 490
special education, 475, 476
specific immune response, 110–111
specificity, 15, 16–17B
spectinamides, 79
spectrin, 507T
spectrin-actin complex, 504
spermatogenesis, 225F, 229B
spermatozoon (sperm), 225F, 228, 229B, 230–232, 453, 476, 556
spherocytosis, 385, 506–508
sphingolipids, 489–490
sphingomyelins, 38
sphygmomanometers, 426, 426F, 427B, 427F, 428
spike proteins, 27, 100F
spinal cord
proton beam therapy, 542
see also central nervous system
spinobulbar muscular atrophy, 468–469
spirochetes, 62, 65, 65F
spirometry, 149, 150B, 150F, 273B
spleen, 136, 317, 415, 416C
erythrocytes, 385
iron metabolism, 385B
LSDs, 490–491
splenectomy, 507–508, 507T
splenectomy, 415
splenomegaly, 170, 186, 401–402
splicing, 401F, 401M
spoiled sweet clover (*Melilotis officinalis*), 394B
spongiform encephalopathies (SEs), 2M, 26, 160T, 462–463B, 462–463T
sprionolactone, 247F
squalene epoxide, 89, 90F
squamous cell carcinomas, 12F, 529, 530F, 552T, 555
SRBCs *see* sheep red blood cells
SRT *see* substrate reduction therapy
SRY mutations, 460F
SSBs *see* single-stranded breaks
SSOP *see* sequence-specific oligonucleotide probe
ssRNA viruses, 28T
ST segment, 429
stable angina, 438–439
staging, 536–537, 549–552, 553, 554F, 554T, 555, 556T
staining, 465, 466, 534–536
Staphylococci, 40–41
Staphylococcus aureus, 37–38, 38F, 39F, 41F, 54–55, 55F
antibody deficiencies, 134
bacteriophages and phagotherapy, 85B
Chédiak–Higashi syndrome, 137
drug resistance, 61B
hyperimmunoglobulin syndromes, 135B
infections of the eyes, 55
infections of the respiratory system, 57
inhibitors of bacterial cell wall biosynthesis, 80
inhibitors of mRNA translation, 78–79
phagocytic defects, 138
sepsis and systemic infections, 65–66
superantigens and toxic shock syndrome, 124B
starvation, 255, 297, 298B, 365, 432
STAT3 *see* signal transducer and activator of transcription 3 protein
statins, 204–205, 441–442, 442F
status asthmaticus, 151
steatorrhoea, 300, 328M, 343, 344F, 346–347, 348, 371
stem cells, 110, 187F, 325, 384F, 562
renewal, 518, 519T
see also haematopoietic stem cell transplantation
stenosis, 437–438, 444
stenting, 439
stercobilin, 318
steroidal anti-inflammatory drugs (SAIDs), 146
steroids, 192, 198–199, 199F
acute promyelocytic leukaemia, 471B
fat-soluble vitamins, 288–289
see also cholesterol; corticosteroids; sex hormones
stethoscopes, 426, 426B, 426F, 427B, 428, 435, 448
stibogluconate, 89F
stomach, 315, 316F
digestion, 322–323
disorders, 337–341
drug receptors, 359
haemosiderin, 385B, 385F
ulcers, 60, 73
vitamin absorption, 331
stomach acids, 47, 60, 322–323, 324B, 336, 339–340
stomach cancer
cigarettes and smoking, 525–526B
general diagnosis, 531
stool acidity tests, 342
STOP codons, 406–407, 505

storage of transplant material, 187
strabismus, 490, 564C
streptococcal toxic shock syndrome (STSS), 41
Streptococci spp., 40–41, 432B
Streptococcus mutans, 58–59, 59F
Streptococcus parvullus, 540
Streptococcus pneumoniae, 55–57, 76F, 134
Streptococcus pyogenes, 37–38, 41, 65–66, 78, 94, 137, 543
Streptococcus sanguis, 58–59
streptogramin, 78
streptokinase, 395
streptokinase-plasmin complex, 349F
Streptomyces hygroscopicus, 184
Streptomyces lincolnensis, 78
Streptomyces peucetius var. *caesius*, 540
Streptomyces pristinaespiralis, 78
Streptomyces spp., 80, 82
Streptomyces tsukubaensis, 184
streptomycin, 77B, 77F
stripping, 447
stroke, 205, 215B, 415, 431, 438, 445, 446, 449, 496F
Strongyloides stercoralis, 47–48
strongyloidiasis, 47–48
strontium (^{89}Sr) radiation therapy, 558
STSS *see* streptococcal toxic shock syndrome
Stuart factor (factor X), 392
Sturli, Adriano, 162
subclinical stages, 3
subcutaneous infection, 125
submetacentric chromosomes, 465F
substance P, 336, 392
substitution mutations, 401–402, 536
substrate reduction therapy (SRT), 491
sugars, 343, 345
  see also carbohydrates
sulphanilamide, 76F
sulphate, 290–291, 291T
sulphonamides/sulpha drugs, 76
sulphonylureas, 214–215, 215F
sulphoraphane, 527F
sulphoxides/sulphones, 88F
Sulston, John, 520B
sunlight, 571B
  see also ultraviolet radiation
superantigens, 41, 65–66, 124B
superbugs, 61B
superoxide dismutases (SODs), 40, 293F, 374–375, 375F
superoxide radicals, 138F, 572
suramin, 86–88
surfactant proteins, 57
SVCT *see* Na^+-coupled vitamin C transporters
swabs for clinical tests, 13–14, 14F
swallowing, 320, 322, 337
sympathetic system, 423, 449
symptomatic poisoning, 364
symptoms and signs, 3, 4T, 69
  breast cancer, 549
  colorectal cancer, 553
  prostate cancer, 557–558
Synacthen, 218–219, 219F
synapsis, 458
syncope, 448
syndrome of inappropriate secretion of ADH (SIADH), 247
syndromes, 4
syngeneic HSCT, 185
syphilis, 65
syringe needles, 14F
systemic circulations, 422
systemic inflammatory response syndrome (SIRS), 65–66
systemic lupus erythematosus (SLE), 129, 134, 140, 144, 146–147
  laboratory tests, 147T
systole, systolic contraction, 422–423, 423F
systolic blood pressure, 426, 428

T

T cell depletion, 186
  bone marrow transplantation, 186–187
T cell receptors (TCR), 118, 122, 122–123F, 185F, 545F
  distinguishing between B and T lymphocytes, 117–118, 119B, 119F
  and immunoglobulin domains, 113F, 113M
  superantigens and toxic shock syndrome, 124B, 124F
T cell-replete HSCT, 136
T lymphocytes, 9, 28, 62B, 102–103, 110F, 117, 123, 129, 384F
  activation of, humoral immunity, 122
  ageing, 570–571
  ALLs, 559–560
  antibody deficiencies, 134
  apoptosis and cancer, 520B
  CAR-T, 544–545
  coeliac disease, 342
  combined T and B lymphocyte deficiencies, 131
  DiGeorge anomaly, 136
  distinguishing between B and T lymphocytes, 118–119B
  epigenetic regulators, targeted therapy, 547
  flow cytometry, 161–162B
  Graves' disease, 142–143
  hyperimmunoglobulin syndromes, 135B
  immunological tolerance, 141B
  immunosuppressants/immunosuppression, 184, 185F
  immunosuppression, 183
  marker proteins, 118T
  SCID, 130, 131–133
  serological HLA typing, 179
  telomeres, telomerase reverse transcriptase, 516B
  type IV (delayed-type) hypersensitivies, 153–154
  see also cytotoxic precursor cells; helper T lymphocyte cells
T waves, 428
TAA *see* tumour-associated antigen
tachycardia, 143, 396, 430, 432
tacrolimus, 182, 183, 184F
tacrolimus–FKBP12 complex, 184
tadalafil, 233F
Taenia solium, 46, 92F
Tagliacozzi, Gaspare, 174B
TAGs *see* triacylglycerols
talazoparib, 546
talimogene, 544
tamoxifen, 542, 551F
Tangier disease, 441M, 483
tapeworms, 46, 60, 92F, 92T, 93, 109
targeted therapy, 545–547
tartar, 337
taste, 316M
Tasuku, Honjo, 544M
taurocholate, 317F
taxol, 456F, 456M
taxol (paclitaxel), 540, 541F
Tay–Sachs disease, 489–490, 489T, 490F
tazemetostat, 547
TB *see* tubercuosis
TBG *see* thyroxine-binding globulin
Tc-99m pertechnetate, 209–210
TCA cycle, 89, 285, 335–336, 368F, 370, 370M, 408
  ethanol toxicity, 370
  G6PD, 408
TCR *see* T cell receptors
teeth, 249, 290–291, 293, 298B, 302B, 320, 320–321B
  canines, 321B
  cementum, 320–321B, 321F
  dental hygiene, 337M
  enamel, 58–59, 297, 301, 302B, 320–321B, 321F
  fluoride deficiency, 301
  incisors, 321B
  number and arrangement, 320B, 320F
  structure, 320–321B, 321F
telocentric chromosomes, 465F
telomerase, 574–575, 575F
telomerase reverse transcriptase (TERT), 516, 516B
telomeres, 516B, 574–575, 575F

telophase, 456, 456F, 458
Temin, Howard, 28M
teniposide, 540
terbinafine, 89, 90F
TERC gene mutations, 575
Teresaki, Paul, 179
Teresaki plates, 179F
terminal ileum, 345
TERT *see* telomerase reverse transcriptase
testes, 224F, 225F
  Leydig cells, 500
  testicular atrophy, 352–353
testicular cancer, 531
  antimicrotubular agents, 540
  hormone therapy, 543
  platinum-based drugs, 541
  topoisomerase inhibitors, 540
testosterone, 192, 193F, 198F, 225–227, 226F, 475–476, 543, 558–559
tetanus toxin (tetanospasmin), 39F
tetracyclines, 55, 78F, 340–341
tetraploid karyoptyes, 476, 477
tetrasomy, 473
tezacaftor, 486
TF *see* tissue factor
TGD gene mutations, 483
TGF *see* transforming growth factor
TH *see* transient hypogammaglobulinaemia
thalassaemias, 385B, 398, 402, 405B, 406–408, 417C
  α thalassaemias, 406–407
  β thalassaemias, 407
  diagnosing and treating, 407–408
thermotherapy, 563B
Thermus aquaticus, 73
THF *see* methylene tetrahydrofolate
thiamine *see* vitamin B_1
thiamine pyrophosphate (TPP), 284, 285F, 304–305, 305F
thiamine transporter proteins (THTRs), 331
thiazide diuretics, 251
thiobarbitone, 373F
thiol-activated lysins, 37–38
third-generation sequencing, 517B
thoracic tumours, 534
threadworms, 47–48
threonine kinase, 9F
thrombin, 392–393, 393F, 394B, 411–412
  see also blood clotting
thrombocytopenia, 103M, 412, 414–415
  diagnosing and treating, 415
thrombocytopenic purpura (TPP), 415
thrombocytosis, 415
thromboembolism, 435–436B
thrombolysis, 439, 439M
thrombolytic therapy, 443–444
thromboplastin, 390T, 392F
  see also tissue factor
thrombosis, 439, 525B
thromboxane A_2, 436B
thromboxanes, 281, 436B
thrombus, 438
  see also blood clotting; plaques; platelets
thrush, 60
thryroiditis/Hoshimoto thyroiditis, 142, 143F
THTRs *see* thiamine transporter proteins
thymidine kinase (TK), 120B
thymidylate kinase, 54B
thymidylate synthase, 399, 538, 539F
thymine (T), 71–74, 402, 454F, 455F
  see also deoxyribonucleic acid
thyroid
  abnormal heart rhythms, 430–431
  cancer, 540, 542
  function tests, 209T, 234C
  goitre, 107–108, 142, 155C, 207F, 209–210, 301, 302F
  hormone disorders, 193, 198–199, 206–210, 206F, 293, 398
  macrocytic anaemia, 399
  microcytic anaemia, 398
  myocarditis/viral cardiomyopathy, 436
  overactive thyroid, 433
  underactive thyroid, 430
thyroidectomy, 210
thyroid-stimulating hormone (TSH), 142–143, 192, 196, 200F, 201F, 204, 207–210, 311C, 449C
thyrotoxicosis, 207F
thyrotrophin-releasing hormone (TRH), 201F, 202–203, 206–207, 207F, 209F
thyroxine (T_4), 206–207, 207F, 208–210
thyroxine-binding globulin (TBG), 206, 208–209
ticks (*Ixodes ricinus*), 48F
tidal volume (TV), 150B
tinidiazole, 89F
tissue biopsy, 536
tissue factor (TF), 392F, 412
tissue necrosis, 348
tissue plasminogen activator (t-PA), 439M
tissue tropism, 28
tissue-specific proteins, 503
tissue-type plasminogen activator (t-PA), 395
tixagevimab, 121B
TK *see* thymidine kinase
TKIs *see* tyrosine kinase inhibitors
TLC *see* total lung capacity
TME *see* tumour microenvironment
TNF *see* tumour necrosis factor
TNM system, 536–537, 553, 554F, 554T, 555, 556T
  breast cancer, 550T
  colorectal cancer, 553, 554F, 554T
  prostate cancer, 557–558
    treatment, 558
  retinoblastoma, 563B
TNT *see* trinitrotoluene
tobacco *see* smoking
toga viruses, 56
tolerance, 61B
  barbiturates, 373
Tolypocladium inflatum gams, 184
tomatoes, 52, 527
tongue, 320
topoisomerase IV-DNA, 76–77, 76F
topoisomerases, 537, 561
  inhibitors, 540
TOR complex-1 (TORC), 184
total lung capacity (TLC), 150B
toxic diseases, 6
toxic mediators, 125
toxic shock syndrome, 124B
toxicity
  antifungal drugs, 91–92
  antibiotics, 76M
  antiprotozoal drugs, 86–88
  fat-soluble vitamins, 300
  hypervitaminoses, 301
  metronidazole, 82
  minerals and trace elements, 332–333
  miscellaneous antibiotics, 82, 83, 85
  ototoxicity, 77B
  superantigens, 41, 65–66, 124B
  thrombocytopenia, 415
toxicology, 357–382
  definitions, 357
  detoxification mechanisms, 360–364
  drug action, distribution and excretion, 357–360
toxins, 36, 37–40
  sepsis and systemic infections, 65–67
  targeted therapy, 545
  vaccines and vaccinations, 69
Toxocara canis, 47, 55
toxoids, 69
Toxoplasma gondii, 44F, 55, 56
toxoplasmosis, 43, 44–45, 85, 86T, 95C
TP53 gene mutations, 518, 525
t-PA *see* tissue plasminogen activator
TPP *see* thiamine pyrophosphate; thrombocytopenic purpura
TR *see* trypanothione reductase
trace elements, 290
  absorption, 332–333
  investigation and laboratory testing, 307
  nutritional disorders, 301–302
  recommended intakes and sources, 291–293, 291T
tracrRNA *see* transactivating CRISPR RNA

transactivating CRISPR RNA (tracrRNA), 548
transcobalamin-vitamin B_{12}, 399F
transcription, 76–77, 82, 198–199, 401F, 454F
transcription factor p53, 9F, 519–520B
transfer RNA (tRNA), 78, 79, 86, 492, 496
transferrin, 333, 334F, 385, 385B
transforming growth factor (TGF), 520B, 525B, 543
transforming retroviruses, 529–530
transfusion and transplantation, 152, 159–190
  ABO blood group system, 162–165
  blood and blood products, 160–162
  blood group systems, 162
  complement, 169–170
  hazards, 170–172
  HSCT, 185–187
  immunosuppressants/immunosuppression, 180–185
  laboratory determinaion of blood groups, 168
  other blood group systems, 166–167
  Rhesus systems, 162, 165–166
  screening, 172–173
  solid tissues and organs, transplantation, 173–177
  transplant definition, 159
transfusion-related acute lung injury (TRALI), 171–172
transglutaminases, 389F, 391, 391F, 393
trans-Golgi, 211
transient hypogammaglobulinaemia (TH), 133–134, 133F
transketolase, 284
transketolase activity, 304–305, 305F, 310C
translation, 77–79, 454F
translocation mutations, 469–470, 469F, 471–472, 471F, 495, 522, 523F, 529
transmembrane domains, 483
transmembrane protein complex, 503F
transplant rejection, 45, 55, 159, 175–176B, 180–181
  see also grafts
transplant surgery, 73
transporter protein PepT1, 328, 329F
transthyretin (TTR), 576–577B
trastuzumab, 550
traumatic diseases, 6
traumatic ulcers, 337
treatment *see* diagnosis and treatment
trehalose, 328
trematodes, 46–47, 92F, 92T
tremelimumab, 544
Treponema pallidum, 62, 65F
TRH *see* thyrotrophin-releasing hormone
triacylglycerols (TAGs), 283, 323, 328F, 330F, 439–441, 440F, 441T, 449C
  see also fatty acids
triazoles, 89, 91
Trichinella spiralis, 47F, 92T
trichinosis, 47
Trichomonas vaginalis, 45F, 65, 66F, 73–74, 89
trichomoniasis, 43, 45, 85, 86T
Trichophyton rubrum, 45F
Trichophyton spp., 45, 89
Trichophyton tonsurans, 45F
tricuspid valve, 421F, 422, 438
tricyclic antidepressants, 435–436B
triiodothyronine (T_3), 206–207, 207F, 208–209
trinitrotoluene (TNT), 358B
trinucleotide repeats, 468–469
triploid karyotypes, 476–477, 476F
triptorelin, 558–559
trisomic zygotes, 476F
trisomy 13 (47,13+), 473–474, 474F
trisomy 18 (14,18+), 474F
trisomy 21, 474–475, 474F, 475F
  see also Down syndrome
tritiated thymidine (^{3}H-TdR), 131B, 132F
tRNA-amino acid complex, 86
troponin, 443, 443B, 444F, 450C
troponin-tropomyosin complex, 443, 443B, 444F
TRPV6 (transient receptor potential cation channel), 332F
trypanomiases, 43, 56, 85, 86T, 88
Trypanosoma brucei, 36B, 43F, 56, 87–88, 87F, 125, 502B
Trypanosoma brucei gambiense, 88
Trypanosoma cruzi, 48, 49–50B, 50F, 87F, 502B
trypanosome infections, 86–88
trypanothione, 86, 87F
trypanothione reductase (TR), 86, 87F
trypsin, 326–327, 331, 535
  see also immunoreactive trypsin
tryptophan, 286F, 344B, 371
tryptophan ethyl ester, 344B
tryptophan loading tests, 305, 306F
tsetse flies (*Glossina* spp.), 48F
TSH *see* thyroid-stimulating hormone
TTR *see* transthyretin
tuberculosis (TB), 251
  Addison disease, 218–219
  Ca^{2+} homeostasis disorders, 251
  cachetin, 531B
  disorders of vitamin nutrition, 299
  drug receptors, 359
  PEM, 294
  type IV (delayed-type) hypersensitivities, 153
tuberculosis (TB), 25–26, 58, 75M
  combination therapy, 75M
  inhibitors of bacterial cell wall biosynthesis, 80
  inhibitors of DNA replication/transcription, 76–77
  miscellaneous antibiotics, 83, 85
tubules, 265–267, 268F, 369
  see also microtubules
tubulin, 539–540
tucatinib, 550
tuft cells, 125
tumour markers, 19C, 531–536, 545
tumour microenvironment (TME), 524, 525B, 528
tumour necrosis factor (TNF), 41, 101–103, 107B, 110, 123, 124T, 139, 144, 178F, 178M, 520B, 531–532B, 531F, 543
tumour protein 53 gene (Tp53), 518, 519–520B, 522B
tumour suppressor genes, 517, 518–521, 529, 547M, 552, 562B
tumour suppressor protein, 471B
tumour-associated antigen (TAA), 544–545, 545F
tumourigenesis, 495, 516
tumours, 3F, 62B, 63T
  chemical carcinogens, 524F
  classification, 537T
  definitions, 514
  of the heart, 436–437, 438
  nucleocytoplasmic transport, 495
  see also cancer
tunics, 424
turmeric plant (*Curcuma longa*), 527F
Turner syndrome, 230, 475
TV *see* tidal volume
T-VEC *see* talimogene laherparepvec
two-hit theory, 521M
Twort, Frederick, 84B
type I hypersensitivities, 148–149
type II hypersensitivies, 152
type III hypersensitivies, 152–153
type IV (delayed-type) hypersensitivies, 153–154
types 1/2 diabetes, 5, 212, 213, 294, 380C, 470, 569, 575, 580–581
typing, 168, 168M, 179–180
  see also karyotyping
tyrosine, 464
tyrosine kinase inhibitors (TKIs), 522, 523F, 562
tyrosine kinases, 545–546

U

ubiquinone, 499
ubiquitin protein ligase activity, 521
UCBT *see* umbilical cord blood transplantation
UDP glucuronyltransferase, 317–318, 349–350

UDP-glucuronate, 362, 364
UICC *see* Union for International Cancer Control
ulcerative colitis, 2, 5, 342, 345
ulcers, 60, 73, 270, 337, 340–341, 436B
  Zollinger–Ellison syndrome, 341
ultrafiltration and reabsorption, 241–242
ultrasound, 7F, 228F, 429F, 534, 555
ultraviolet (UV) radiation, 289F, 521, 528–529, 571B
umami flavour, 316M
umbilical cord blood transplantation (UCBT), 132B, 185
umbilical cord stem cell transplantation, 187
unconjugated bilirubin, 317–318
unconjugated hyperbilirubinaemia, 349–350
underactive thyroid, 430
undernutrition, 303
unfolded protein response (UPR), 487
Union for International Cancer Control (UICC), 536, 555, 556T
United States Food and Drug Adminsitration *see* Food and Drug Adminsitration
unstable angina, 439
u-PA *see* urokinase
UPR *see* unfolded protein response
uraemia, 213
uranium mining and cancer, 555
urate balance disorders, 239, 256–258
  hyperuricaemia, 256, 257–258
  hypouricaemia, 256
  urate structure, 256F
  uric acid in its native state, 256M
urea, 433
urea breath test, 340, 341F
urea transporters, 167
urethritis, 60, 61F
uric acid, 256F, 256M, 257, 301, 302
uricolysis, 256
uridine (U), 64B, 64F, 406–407, 454–455
urinary tract infections (UTIs), 60
  inhibitors of bacterial cell wall biosynthesis, 80
  inhibitors of DNA replication/transcription, 76
  miscellaneous antibiotics, 82
urine, urinary system, 12
  polyuria, 203, 245, 248
  see also kidneys
urobilinogen, 318F
urogenital system, infections of, 60–65
urokinase (u-PA), 349M
ursodiol, 350F
urticaria (skin rashes), 152
uterine cancer, 528, 542
UTIs *see* urinary tract infections
UV *see* ultraviolet radiation

V

vaccines, vaccination, 23, 27F, 29, 31–33, 39, 42B, 67–69
  baby and chilhood programmes, 68B
  BCG, 153, 543
  Edward Jenner, 68–69M
  Haemophilus influenzae, 56M
  HIV and AIDS, 63B
  human papilloma viruses, 531
  immunotherapy, 543, 544
  MMR, 68, 68B, 131
  SCID, 131
  type IV (delayed-type) hypersensitivies, 153
Vaccinia spp., 27
vaccinia virus, 68–69M
vacutainers, 14F
vaginal cancer, 542
vagus nerve, 423F
valacyclovir, 337
valine, 344B
valve cusps, 446–447, 447F
valves of Kerckring, 328
van der Waals forces, 488M
vancomycin, 61B, 80, 81F
vanillylmandelic acid (VMA), 223B
vaping, 526B
vardenafil, 233F
variant angina, 439
variant Creutzfeldt–Jakob disease (vCJD), 26F, 160T, 462–463T, 46203B
variants *see* mutations
Varicella zoster, 29, 29F, 54, 415
  see also chicken pox
varicose veins, 371, 380C, 445, 446–447, 447F
  treatment, 447
vascular disease, 294, 295T
  see also coronary heart disease
vascular endothelial growth factors (VEGFs), 525B, 546, 547
vasoactive intestinal peptide (VIP), 336
vasoconstriction, 439
vasodilatory drugs, 433–434
vasopressin, 196, 202F, 243F, 412–413
  see also antidiuretic hormone; arginine vasopressin
VC *see* vital capacity
vCJD *see* variant Creutzfeldt–Jakob disease
vectors, 48F, 85
vegetation, 430
VEGF *see* vascular endothelial growth factor
VEGFs *see* vascular endothelial growth factors
veins, 383, 422, 424–426, 424F
venetoclax, 520B, 520F, 546
venous thromboembolism, 435–436B
ventricles, 421F, 422–423, 429, 431–432, 437–438
  fibrillation, 432
  hypertrophic cardiomyopathy, 436–437
  see also heart; tachycardia
venules, 424F
vermifuges/vermicids, 92
vertebrates, 48–50
vertigo, 77B
very long-chain fatty acids (VLCFAs), 499–500, 502, 509C
very low-density lipoproteins (VLDLs), 283, 440–441, 440F, 440T, 441
vestigial caecum, 336F
Vibrio cholerae, 36, 60
Victoria I, 414B
villi, 325F, 342F
  see also chorionic villus sampling
vinblastine, 540F
vinca alkaloids, 540F
vincristine, 540F, 563B
vinorelbine, 540
vinyl chroride, 524T
VIP *see* vasoactive intestinal peptide
viral cardiomyopathy, 436
viral meningitis, 56
viral replication cycle, 28, 29F
virilism, 230
virions, 27F, 30F, 32F, 33F, 100F, 111F, 529F
virulence factors, 24
viruses, viral infections, 27–33
  and cancer, 529–531
  cardiac failure, 433
  Chédiak–Higashi syndrome, 138
  DiGeorge anomaly, 136
  drug resistance, 61B
  endoplasmic reticulum, disorders of, 487
  epidemics and pandemics, 30–32
  helper T lymphocyte activation, 123
  immunodeficiency diseases, 129
  NGS methods, 517B
  nomenclature and classification, 27–28
  nucleocytoplasmic transport, 495
  oncogenes, 517
  thrombocytopenia, 414–415
  viral infections, 28–30
  see also antiviral drugs; immune system, immune response
vital capacity (VC), 150B
vitamers, 284, 286–290, 331
vitamin A, 5–6, 284F, 284T, 287–289, 293, 300F, 301, 332
vitamin B deficiency, 528
vitamin B_1 (thiamine), 284T, 285F, 298B, 304–305, 331
vitamin B_2 (riboflavin), 285F, 299, 304–305, 331

vitamin B_3 (niacin or nicotinic acid), 286F, 299F, 301, 305, 331, 344B, 371
vitamin B_5 (pantothenic acid), 282F, 286, 287F, 299, 331F
vitamin B_6 (pyridoxine, pyridoxal, pyridoxamine), 286F, 287F, 299, 301, 305, 306F, 331
vitamin B_7 (biotin, H), 286, 287F, 299F, 305, 331F
vitamin B_9 (folate, folic acid), 507–508, 528, 537, 538F
vitamin B_9 (folic acid, folate), 182, 286–287, 288F, 299–300, 305, 331, 343, 371, 399, 439
vitamin B_{12} (cobalamin), 286–287, 288F, 300, 305–306, 323F, 331F, 339–340, 343, 399F, 439
vitamin C (ascorbic acid), 284F, 284T, 287, 288F, 290F, 300, 306, 331, 340, 572
vitamin D, 100M, 250–252, 284T, 289, 293, 300, 301, 307, 311C, 332, 343, 527
vitamin D_3 (cholecalciferol), 289F
$1,25(OH)_2D_3$ form of vitamin D, 332
vitamin E (α-tocopherol), 284T, 287, 289–290, 289–290F, 294, 300, 307, 332, 340, 572, 578–579
vitamin H *see* vitamin B_7
vitamin K, 284T, 289, 290F, 300–301, 301F, 301M, 307, 332, 343, 394B, 411, 415, 435B
vitamins, 284–290, 371, 499
  absorption, 331–332
  hormone classification, 192
  investigation and laboratory testing, 304–307
  nutritional disorders, 297–301
  see also fat-soluble vitamins; water-soluble vitamins
VLCFAs *see* very long-chain fatty acids
VLDL *see* very low-density lipoprotein
VLDLs *see* very low-density lipoprotein
VMA *see* vanillylmandelic acid
vomiting, 324B
von Decastello, Alfred, 162
von Willebrand disease (vWD), 412–413
  diagnosing and treating, 412–413
von Willebrand factor (vWF), 391, 392, 412–413
voriconazole, 91
vorinostat, 547
vWD *see* von Willebrand disease
vWF *see* von Willebrand factor

W

waist-to-height ratio (WHtR), 295B, 304
warfarin, 283, 300–301, 301F, 301M, 394B, 394F, 435–436B
WAS-related disorders, 136–137
waste products, 571, 574
WAT *see* white adipose tissue
water, 239, 244–245
  absorption, 333–335
  balance, 283
  and diet, 283
  distribution in the body, 244F
  mouth, 320
  osmolality, 245B
water treatment (purification), 23, 24F
water-soluble vitamins, 284–287, 304–306
  absorption, 332
  clinical disorders from deficiencies, 298–300
  investigation and laboratory testing, 304–306
'wear and tear' theories, 571–572
weight management, 235C, 310C, 434, 458
  see also body mass index; obesity
Wernicke–Korsakoff syndrome, 298, 298B
wet form beriberi, 298
white adipose tissue (WAT), 531–532B
white blood cells (WBCs) *see* leukocytes
WHO *see* World Health Organization
whole rhizomes, 527F
whooping cough, 67–68
WHtR *see* waist-to-height ratio
Wiener, Alexander, 165
wild-type genes, 517
  oncogenes, 517
  see also proto-oncogenes
Williams, Cicely, 294
Wilms tumour
  antimicrotubular agents, 540
  topoisomerase inhibitors, 540
Wilson disease, 292B
wingless-related integration site (^{a}WNT), 518, 519T, 553
Wiskott–Aldrich syndrome, 136–137, 461T
WNT/β-catenin pathways, 522
World Cancer Research Fund, 526
World Health Organization (WHO), 1–2, 31, 32, 33, 35, 35B, 514, 528, 561
  baby and chilhood vaccination programmes, 68B
  Chagas disease, 49B, 49T
  Edward Jenner, 68–69M
  Haemophilus influenzae, 56M
  helminths, 46
  HIV and AIDS, 62B
  infections of the respiratory system, 57–58
  inhibitors of bacterial cell wall biosynthesis, 80
  malaria, 42B
  trypanomiases, 43
World War I, 31, 173, 358B
World War II, 173, 525B
worms *see* helminths
Wuchereria bancrofti, 47, 92T, 93

X

X chromosomes
  haemophilia A, 413, 413M
  sex-linked genetic diseases, 450–451, 453F, 460–461, 460F
xanthase oxidase, 302
xanthine oxidase, urate balance disorders, 256F
xanthomatas, 441, 450C
xanthurenate, 305, 306F
xenobiotics, 357, 358B, 360, 362F, 408M
xenogeneic grafts, 175
xenogeneic transplantation, 175–176B
XenoMouse transgenic mouse models, 121B
Xenopus laevis (toad spp.), 492
xeroderma pigmentosum, 528–529
xerostomia, 337
X-irradiation *in utero*, 528
XLA *see* X-linked agammaglobulinaemia
X-linked agammaglobulinaemia (XLA), 134–135, 136, 461T
X-linked congenital neutropenia (XLN), 136
X-linked disorders, 414, 459, 460–461, 461T
  complement deficiencies, 140
  Leber hereditary optic neuropathy, 497–498
  muscular dystrophies, 505
  SCID, 131–132
X-linked thrombocytopenia (XLT), 136–137
XP gene mutations, 528–529
Xp28, *ABCD1* mutations, 499–500
X-rays, 3F, 7, 7M, 27F, 429F, 433, 529, 541
  angiography, 430
  arthritis, 579
  COPD, 273B, 273F
  haematomas, 417C
  hiatus hernias, 338
  ischaemic heart disease, 434
  lung cancer, 555
  mammography, 522B
  peritonitis, 320B
  pituitary function, 202
  proton beam therapy, 542
  tumour detection/localisation, 533–534
XX chromosomes, 453, 460–461, 465, 466F, 467F, 468–469, 468B, 473–476
  muscular dystrophies, 505
XY chromosomes, 453F, 460–461, 460F, 473–476
  Menkes disease, 292B

xylene, 534
xylose, 343, 345

Y

Y chromosomes *see* XY chromosomes
YCT529, 229B
yeasts (thrush), 337
Yersinia enterocolitica, 39
Yersinia pestis, 34–35B, 36, 48
Yersiniae, 36–37

Z

ZAG *see* zinc α2-glycoprotein
Zellweger, Hans, 502
Zellweger spectrum (ZS) disorder, 500, 502
zinc α2-glycoprotein (ZAG), 531–532B
zinc (Zn^{2+}), 39F, 132F, 291T, 302F, 307, 333, 371, 388F
Zollinger–Ellison syndrome, 337, 341
zoonoses, 24M
ZS *see* Zellweger spectrum disorder
zygotes, 228F, 473, 476F, 514M
zymogens, 325M, 327T, 390–391, 439M